

# 开放数据与数据安全的政策协同度研究<sup>\*</sup>

马海群<sup>1</sup> 张涛<sup>2</sup> 易扬<sup>3</sup>

(1 黑龙江大学信息资源管理研究中心, 哈尔滨 150080;

2 黑龙江大学信息与网络中心, 哈尔滨 150080;

3 黑龙江大学数学科学学院, 哈尔滨 150080)

**摘要:** [目的/意义] 长期以来, 数据的开放与保护、共享与安全是数据政策制定与实施过程中所面临的严峻挑战, 探讨数据政策中开放数据和数据安全的协同发展对政府优化信息资源配置、促进大数据产业创新有着重要意义。[方法/过程] 文章首先通过政策文本计算的方法对开放数据子系统及数据安全子系统有序度进行计算, 随后构建两者复合系统协同度模型, 最后分析了数据政策中的开放数据与数据安全之间的相互作用, 并阐述了两者的协同发展的机理。[结果/结论] 对所收集的 446 条数据政策分析可知: 我国数据政策中开放数据和数据安全子系统间良性协同发展机制尚未形成, 但在干政策和枝政策中开放数据和数据安全主题间有序度和协同度发展趋势较好, 这也说明了在未来通过政策调整可能会使开放数据和数据安全有较好的协同关系, 最终提出了促进该复合系统协同发展的对策建议。

**关键词:** 开放数据 数据安全 政策协同 政策文本计算

**分类号:** D63; TP309.2

**DOI:** 10.31193/SSAP.J.ISSN.2096-6695.2020.01.03

## 0 引言

随着数字化经济时代的到来, 大数据将成为推动经济社会快速发展的关键因素, 数据引发的变革正在从商业领域向社会治理和个人生活领域全面拓进, 数据正在改变所有组成文明的要素。数据的开放与保护、共享与安全是信息化社会健康发展的重要力量, 它不仅需要政策层面的激励, 还需要技术层面的支撑。由于技术特性和法律特性的限定, 在开放数据的定义中, 应当使用

<sup>\*</sup> 本文系国家自然科学基金重点项目“开放数据与数据安全的政策协同研究”(项目编号: 15ATQ008)和黑龙江省哲学社会科学规划项目“智能+视阈下基于语料库的数据政策模型构建与实证研究”(项目编号: 19TBQ073)的成果之一。

[作者简介] 马海群 (ORCID: 0000-0002-2091-7620), 男, 教授, 博士生导师, 博士, 研究方向为信息政策与法律, Email: mahaiqun@sina.com.cn; 张涛 (ORCID: 0000-0002-3367-4541), 男, 博士研究生, 高级工程师, 研究方向为文本计算、数据库研究, Email: zhangtao@hlju.edu.cn; 易扬 (ORCID: 0000-0001-5496-644X), 女, 硕士研究生, 研究方向为李代数, Email: 1228231579@qq.com。

涵盖敏感信息、安全情报、个人隐私保护的潜台词来有效确保数据安全,并体现开放数据与数据安全协同性的内涵。在大数据时代,政府应兼顾数据的“开放”和“安全”,过度提高政府数据开放的深度和广度会引发第三方对数据的不当和不法利用,但过多聚焦于政府数据安全及隐私保护会阻碍对数字资源的合理使用,降低政府管理的透明度和公民对政府的满意度,因此两者有机结合并协同发展就变得尤为重要。

近年来学界就开放数据与数据安全政策间的关联关系和影响因素进行了一系列研究,如:孙瑞英等人从太极图哲理内涵的角度提出了开放数据与数据安全政策协同的客观要求,并通过统一论视角、和谐论视角、运动观视角阐释了开放数据与数据安全政策协同关系<sup>[1]</sup>;毛子骏等人将我国12个省市出台的与政府数据开放相关的政策内容分为政策目标与政策措施,从纵向政策协同、横向政策协同的维度对政策间差异进行分析<sup>[2]</sup>;王德庄等人从利益相关者理论视角对科学数据开放政策与个人数据保护政策的政策协同进行研究<sup>[3]</sup>;丁红发等人从面向数据生命周期视角对政府数据开放的数据安全与隐私保护进行分析<sup>[4]</sup>;张彬等人运用文献梳理、对比研究等方法对数据政策中数据开放和安全隐私做了总结分析,并对我国数据治理提出了相关政策建议<sup>[5]</sup>。

综上所述,虽有部分学者对开放数据和数据安全政策协同关系做出研究,但目前尚无学者从定量的角度对开放数据和数据安全政策协同度进行研究。本文运用协同学理论和政策文本计算的方法对开放数据与数据安全子系统之间的协同发展机理进行分析,并构建复合系统协同度测度模型,通过实证研究客观分析两者间发展演化趋势,从而为政府制定数据政策提供科学化的理论支撑和决策参考。

## 1 协同发展机理

本文借鉴孟庆松等人<sup>[6]</sup>的研究成果,将开放数据和数据安全视作复合系统的两个子系统,基于对两个子系统的有序度和协同度计算,给出开放数据与数据安全政策协同的对策及建议。

协同学是一门跨越社会科学与自然科学的横断学科<sup>[7]</sup>,系统之间在发展演化过程中彼此和谐一致是协同的重要内容,为实现上述和谐一致需要对系统采取一系列调控活动,所有可能的调控活动及其所遵循的程序与规则称为协同机制<sup>[8]</sup>。开放数据与数据安全之间存在繁复的非线性关系(如图1所示),一方面因为数据开放方式、数据政策引导等因素的局限性而引发数据安全问题;另一方面由于过度强调数据安全,而阻碍数据开放共享的进程,由此影响了产业发展的进程<sup>[9]</sup>。开放数据子系统在数据共享交换及数据授权使用过程中,通过数据安全子系统确定数据保护边界、数据传输加密等手段为开放数据提供支持<sup>[10]</sup>。开放数据与数据安全既相互依存,又有不同的价值导向,只有实现协调有序,才能解决两者之间矛盾冲突所带来的问题。

## 2 协同度模型构建

本文将通过政策文本计算的方式计算以开放数据和数据安全为主题词的协同度数值,设置 $n_1$ 为开放数据主题词数, $n_2$ 为数据安全主题词数, $m$ 为语料总数。开放数据与数据安全复合系

统由两个子系统构成, 分别表示为  $S_1$  (开放数据子系统) 和  $S_2$  (数据安全子系统)。具体实现步骤如下:

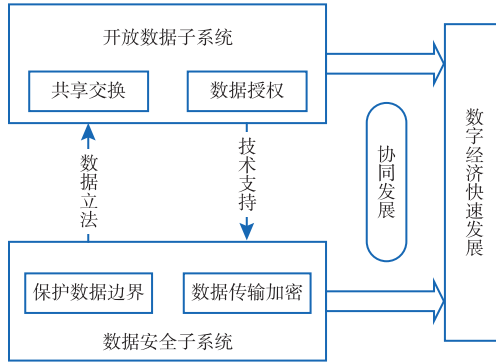


图1 开放数据与数据安全相互作用机理

第一步, 根据开放数据和数据安全两个主题构建矩阵  $S_1$  和  $S_2$ , 形成两个子系统统一由  $S_{\kappa}, \kappa \in [1, 2]$  表示, 设开放数据和数据安全的系统发展过程中的序参量具体如下:

$$S_1 = \{x_{11}, x_{12}, \dots, x_{1i}\}, i = (1, 2, \dots, n_1), n \geq 1, \beta_{1i} \leq x_{1i} \leq \alpha_{1i}, i \in [1, n_1]$$

$$S_2 = \{y_{21}, y_{22}, \dots, y_{2j}\}, j = (1, 2, \dots, n_2), n \geq 1, \beta_{2j} \leq y_{2j} \leq \alpha_{2j}, j \in [1, n_2]$$

第二步, 通过对  $S_1$ 、 $S_2$  进行标准化来消除原始数据不同量纲的影响, 先取均值再做标准差:

$$S'_i = \frac{x_{1i} - \bar{x}}{X_i} \quad (1)$$

在公式 (1) 中:  $S'_i$  为标准化数据;  $\bar{x}$  表示变量  $x_{1i}$  的均值;  $X_i$  表示变量  $x_{1i}$  的标准差  $X_i = \sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} (x_{1i} - \bar{x})^2}$ 。

第三步, 运用相关矩阵赋权法确定各指标权重。步骤如下: 设指标体系中包含  $n_1$  个指标, 它们的相关矩阵  $A$  为:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n_1} \\ a_{21} & a_{22} & \dots & a_{2n_1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_1 1} & a_{n_1 2} & \dots & a_{n_1 n_1} \end{bmatrix} \quad (2)$$

其中  $a_{ii} = 1, A_i = \sum_{j=1}^{n_1} |a_{ij}| - 1, i = 1, 2, \dots, n_1, j = 1, 2, \dots, n_1$ ,  $A_i$  表示第  $i$  个指标对其他  $(n_1 - 1)$  个指标的总影响。  $A_i$  越大, 表明第  $i$  个指标在整个指标体系中的重要性越高, 故赋予的权重值应越大, 因此将  $A_i$  归一化处理可得到相应各指标的权重, 具体如公式 (3):

$$\lambda_i = \frac{A_i}{\sum_{i=1}^{n_1} A_i} \quad (3)$$

按上述步骤计算协同度测度指标体系中各二级指标权重,由于本文中协同度测度模型计算过程中不需要一级指标权重,故未计算。

第四步,计算开放数据子系统的有序度。假定  $x_{11}, x_{12}, \dots, x_{1k}$  是刻画开放子系统的正向指标,即取值越大,开放数据子系统有序度越高。 $x_{1k+1}, x_{1k+2}, \dots, x_{1n}$  为负向指标,即取值越小,开放数据子系统有序度越低。

$$\text{则有 } u_1(x_{\kappa i}) = \begin{cases} \frac{x_{\kappa i} - \beta_{\kappa i}}{\alpha_{\kappa i} - \beta_{\kappa i}}, & i \in [1, k] \\ \frac{\alpha_{\kappa i} - x_{\kappa i}}{\alpha_{\kappa i} - \beta_{\kappa i}}, & i \in [k+1, n] \end{cases}, \kappa \in [1, 2]$$

根据上面公式可算出开放数据序参量的有序度。

由上述定义知  $u_1(x_{i_i}) \in [0, 1]$ ,  $u_1(x_{i_i})$  越大,  $x_{i_i}$  对开放数据子系统的有序度越大。从总体上看,  $x_{i_i}$  对开放数据子系统的有序度的贡献可以通过  $u_1(x_{i_i})$  来实现,即:

$$u_1(S_{i_i}) = \sum_{i=1}^{n_i} \lambda_i u_1(x_{i_i}), \lambda_i \geq 0, \sum_{i=1}^{n_i} \lambda_i = 1,$$

则称  $u_1$  为开放数据子系统的有序度。

第五步,重复上述过程可以利用  $u_2(S_{2_j}) = \sum_{j=1}^{n_2} \lambda_j u_2(y_{2_j}), \lambda_j \geq 0, \sum_{j=1}^{n_2} \lambda_j = 1$  计算出数据安全子系统的有序度。

第六步,通过两个系统有序度数值来计算复合系统协同度。设初始时刻为  $t_0$ ,开放数据和数据安全子系统的有序度分别为  $u_1^0(S)$  和  $u_2^0(S)$ 。在复合系统发展演变过程中的另一时刻为  $t_1$ ,开放数据和数据安全子系统有序度分别为  $u_1^1(S)$  和  $u_2^1(S)$ ,设复合系统的协同度为  $C$ ,如下定义:

$$C = \text{sig}(\cdot) \times \sqrt{(|u_1^1(S) - u_1^0(S)| \times |u_2^1(S) - u_2^0(S)|)} \quad (4)$$

$$\text{其中 } \text{sig}(\cdot) = \begin{cases} 1, & u_1^1(S) - u_1^0(S) > 0, u_2^1(S) - u_2^0(S) > 0 \\ -1, & \text{其他} \end{cases}$$

由上述公式可知,协同度  $C$  为正值且必须满足一个条件,即开放数据和数据安全子系统在  $t_0$  时刻的有序度均小于其在  $t_1$  时刻的有序度。当这两个子系统中任何一个在  $t_0$  时刻的有序度大于其在  $t_1$  时刻的有序度时,复合系统协同度  $C$  就会呈现负值。若只有一个子系统数值较大,则复合系统协同度并不会出现同等程度的提升,因此两个子系统的发展状况才是复合系统协同度模型测度的主要依据,与两个子系统的有序度都有密切关联<sup>[11]</sup>。由公式(4)可知,  $C \in [-1, 1]$ ,协同度数值越高,表明协同效应越好,具体等级划分参照夏业领<sup>[12]</sup>的研究(如表1所示),最后通过计算数值来判定两者间的协同关系。

表1 复合系统协同度评价指标

协同度	协同度等级
$-1 \leq C < 0$	严重不协同
$0 \leq C < 0.4$	不协同

续表

协同度	协同度等级
$0.4 \leq C < 0.6$	轻度不协同
$0.6 \leq C < 0.8$	基本协同
$0.8 \leq C < 0.9$	良好协同
$0.9 \leq C < 1$	优质协同

### 3 实证研究

#### 3.1 数据来源

本文选取自建语料库<sup>[13]</sup>中 446 条数据类政策文本, 并将其分为根政策、干政策和枝政策三个层级<sup>[14]</sup>。根政策包括全国人大和国务院颁布的 19 条政策法规; 干政策包括国家各部委颁布的 41 条政策; 枝政策包括各省、市、区等地方政府所颁布的 386 条政策。

利用 NLPPIR-Parser 对数据政策文件进行关键词提取, 通过分词、去停用词等操作, 提取政策文本中名词、形容词、动词等核心词性的词语, 再利用 LDA 进行词语聚类分析<sup>[15]</sup>, 人工筛选出和开放数据与数据安全相关的关键词 46 个, 把相关关键词按照 Rothwell 和 Zegvold<sup>[16]</sup>所提出的供给型、需求型和环境型三类政策工具进行手动划分, 在此环节中, 采用德尔菲专家调查方法对关键词进行手动分类(如表 2 所示)。

表 2 开放数据和数据安全提取关键词表

政策措施 政策目标	供给型	需求型	环境型
开放数据	共享交换、共享使用、共享开放、互联互通、范围边界、共享共用	信息共享、数据共享、信息公开、政务公开、数据开放、条件共享	数据管理、共享平台、开放平台、制度对接、共享评估
数据安全	应急处置、应急预案、突发事件、资源配置、人才培养、国家秘密、商业秘密、数据保密	技术研发、安全保护、职责明确、等级保护、安全保障、权益保护	知识产权、数据立法、舆情监控、风险评估、监督管理、监督检查、监测预警、隐私安全、隐私保护、数据伪造

指标体系是反映系统之间协同发展的重要元素, 构建科学合理的指标体系能够使协同度计算更加精准, 根据以上政策工具的分析, 把供给型、需求型、环境型作为开放数据子系统和数据安全子系统的序参量, 把所提取的关键词作为其二级指标构建具体指标体系(如表 3 所示)。

#### 3.2 实证分析

本文利用政策文本计算的方法分析 2008 ~ 2018 年的 446 条数据政策中开放数据与数据安全主题协同发展关系, 通过构建基于序参量的复合系统协同发展模型<sup>[17]</sup>并计算出有序度和协同度数值, 从而对开放数据与数据安全主题间协同发展进行实证研究。

表3 开放数据子系统和数据安全子系统协同度指标体系

子系统	序参量	二级指标
$S_1$ 开放数据子系统	$S_{11}$ 供给型	$x_{11}$ 共享交换
	$S_{12}$ 需求型	$x_{12}$ 信息共享
		$x_{13}$ 数据共享
	$S_{13}$ 环境型	$x_{14}$ 共享平台
		$x_{15}$ 开放平台
$S_2$ 数据安全子系统	$S_{21}$ 供给型	$y_{21}$ 应急处置
		$y_{22}$ 国家秘密
	$S_{22}$ 需求型	$y_{23}$ 权益保护
		$y_{24}$ 安全保护
	$S_{23}$ 环境型	$y_{25}$ 知识产权
		$y_{26}$ 数据立法

### 3.2.1 根政策

通过对 2008 ~ 2018 年全国人大及国务院发布的 19 条数据类根政策进行实证研究,并分别计算开放数据和数据安全子系统的有序度(如图 2 所示)和复合系统协同度(如图 3 所示),分析发现:开放数据子系统和数据安全子系统的有序度数值在  $[0.2, 0.5]$  区间震荡,开放数据子系统有序度波动较大。由于国家在 2008 年和 2010 年分别发布《中华人民共和国政府信息公开条例》和《国务院办公厅关于做好政府信息依申请公开工作的意见》等政策文件,这也是我国关于开放数据问题较权威的政策文件,因此在 2008 年和 2010 年开放数据子系统有序度为最高的而数据安全子系统有序度波动较小,且上升趋势不是很显著。在根政策中开放数据子系统与数据安全子系统的协同度数值在  $[-0.15, 0.05]$  区间震荡,虽然整体数值较低,但在 2015 年出现峰值,这是因为 2015 年是我国大数据发展的关键年,国务院接连印发《促进大数据发展行动纲要》、《关于运用大数据加强对市场主体服务和监管的若干意见》、《关于加快构建大众创业

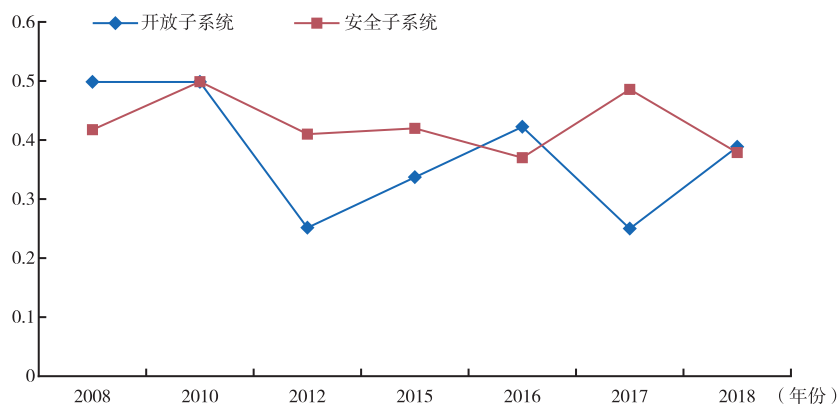


图2 根政策中开放数据子系统与数据安全子系统有序度发展趋势



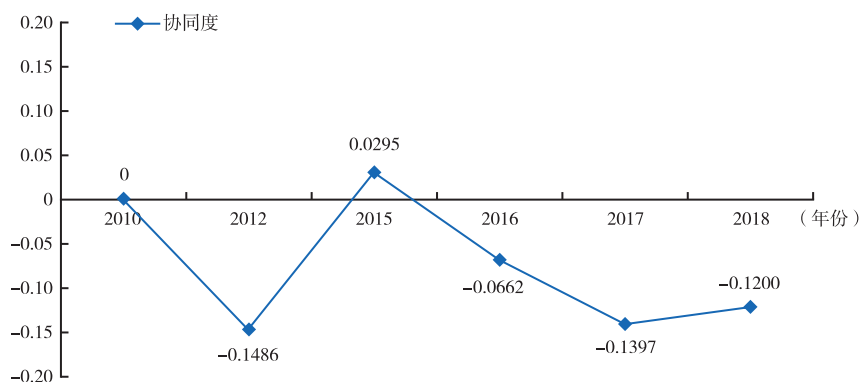


图3 根政策中开放数据子系统与数据安全子系统协同度发展趋势

万众创新支撑平台的指导意见》等重要的政策文件，在这些政策文件中对开放数据和数据安全都有比较明确的阐述，因此该年复合系统协同度数值达到最高为 0.0295，但随后三年出现了大幅度下降，这表明在根政策中开放数据与数据安全良性协同发展机制尚未形成。

### 3.2.2 干政策

通过对 2007 ~ 2018 年国家各部委发布的 41 条数据类干政策进行实证研究，并分别计算开放数据和数据安全子系统间有序度（如图 4 所示）和复合系统协同度（如图 5 所示），分析发现：开放数据子系统和数据安全子系统的有序度在  $[0, 0.5]$  区间震荡，开放数据子系统有序度波动较大；其中 2007 年数据安全子系统的有序度为 0，这是由于 2007 年由商务部办公厅发布关于落实《政府信息公开条例》有关事项的通知中只提及了信息公开、信息开放等相关内容，并未对数据安全问题做明确规范，因此数据偏差较大；2009 ~ 2018 年数据安全子系统有序度波动较小，相对较为稳定，且上升趋势不是很显著。在干政策中，开放数据与数据安全子系统的协同度数值在  $[-0.15, 0.05]$  区间震荡，整体数值较低，震荡幅度较大。在 2018 年 7 月 12 日，中国信息通信研究院在中国互联网大会上发布了《大数据安全白皮书（2018 年）》，白皮书从大数据开放利用角度出发梳理了数据安全威胁及相应的数据安全保障技术的发展情况，该文件为大数据产业和安全技术发展提供依据和参考。因此复合系统协同度数值从 2017 年的低谷上升到 2018 年的 0.1193，虽然 2018 年数值有所提升，从整体来看，干政策中开放数据与数据安全良性协同发展机制也未形成。

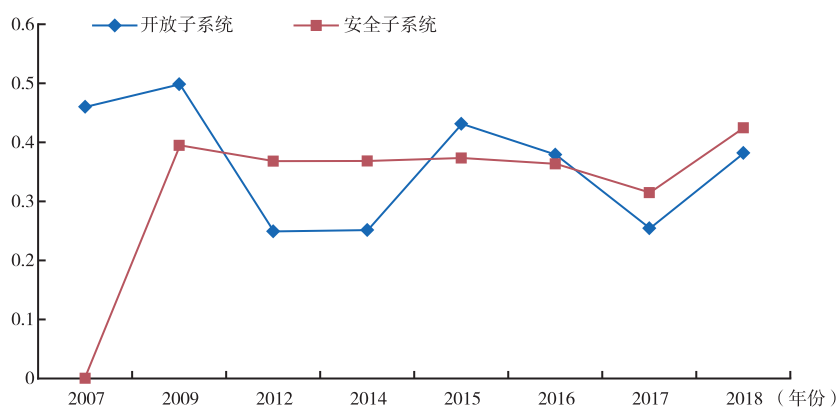


图4 干政策中开放数据子系统与数据安全子系统有序度发展趋势

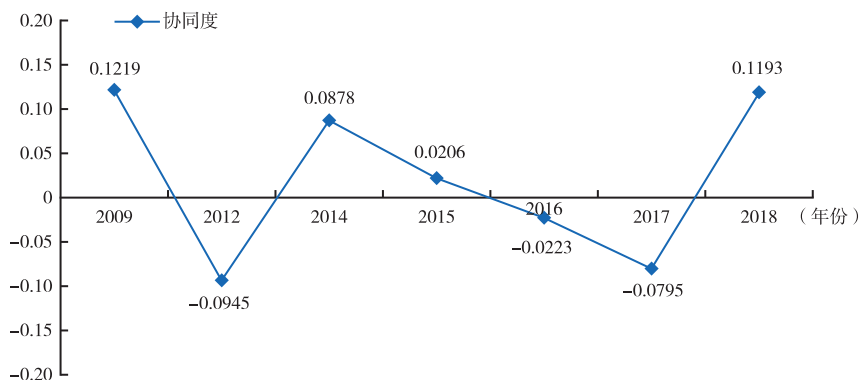


图 5 干政策中开放数据子系统与数据安全子系统协同度发展趋势

### 3.2.3 枝政策

通过对 2008 ~ 2018 年各省、区、市发布的 386 条数据类枝政策进行实证研究，并分别计算开放数据和数据安全子系统的有序度（如图 6 所示）和复合系统协同度（如图 7 所示），分析发现：开放数据子系统和数据安全子系统的有序度在  $[0, 0.5]$  区间震荡；在 2008 ~ 2010 年开放数据子系统和数据安全子系统的有序度波动较大，由于这三年数据类政策发布总计为 4 条，故参考意义不大；

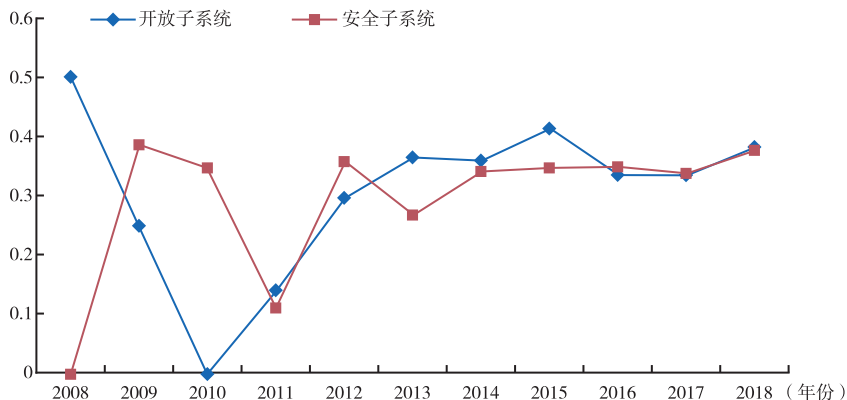


图 6 枝政策中开放数据子系统与数据安全子系统有序度发展趋势

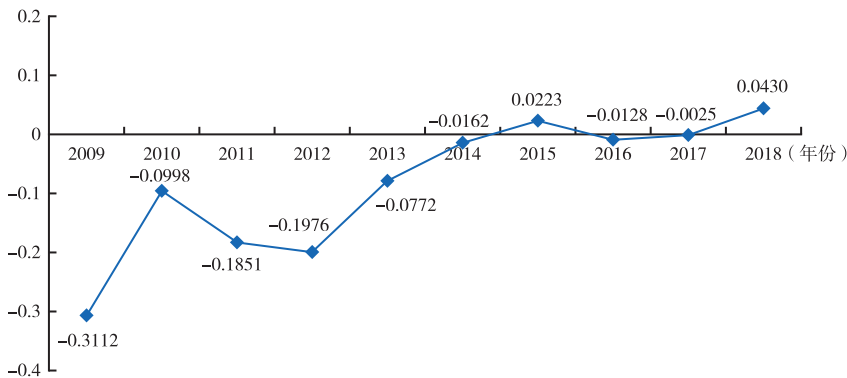


图 7 枝政策中开放数据子系统与数据安全子系统协同度发展趋势



从 2011 年起到 2018 年, 枝政策数量处于上升趋势, 开放数据子系统和数据安全子系统的有序度相对稳定, 并且曲线类似, 数值在  $[0.2, 0.4]$  这个区间范围波动, 只有 2015 年开放数据子系统有序度略高于 0.4, 但是政策有序趋势相对较为平稳。在枝政策中开放数据子系统与数据安全子系统的协同度数值在  $[-0.3, 0.1]$  区间震荡, 整体数值较低, 但是从 2012 年起协同度数值就呈现上升趋势, 并在 2018 年达到最高的 0.0430, 数据表明: 在枝政策中各省、区、市政府落实国家层面政策时, 步调较为一致, 虽然开放数据与数据安全协同度数值较低, 但是从 2012 年起开放数据子系统和数据安全子系统的有序度关联性较好, 并且从 2011 年起复合系统协同度也处于稳步上升的阶段, 但该趋势能否持续还需要持续关注后续的发展情况。

## 4 对策建议

通过对 2008 ~ 2018 年 446 条数据政策复合系统协同度的分析研究表明: 在根政策和干政策中开放数据和数据安全子系统的有序度上升趋势不明显, 复合系统协同度波动较大, 说明子系统间良性协同发展机制尚未形成, 但干政策中有序度关联性较好, 说明在干政策中两个子系统步调较为一致, 未来能否持续需关注后续政策。在枝政策中开放数据和数据安全子系统的有序度关联性较好, 这说明各省、区、市政府落实国家层面政策时, 两个子系统有着较好的一致性。在枝政策中开放数据与数据安全协同度数值虽然较低, 但从 2011 年起复合系统协同度处于稳步上升的趋势。总体来说, 虽然开放数据和数据安全子系统间良性协同发展机制尚未形成, 但是干政策、枝政策中主题间有序度和协同度发展趋势较好, 也说明了在未来通过政策的进一步调整可能会使开放数据和数据安全有较好的协同关系。

结合以上数据实证分析, 提出如下对策建议。

### 4.1 构建统一政府数据开放共享平台

政府要从构建统一政府数据开放共享平台的角度着手提升开放数据在政策文件中的权重, 该平台不但要用于政府数据的统一汇聚、安全存储、开放共享, 还要明确数据标准及数据所开放的领域<sup>[18]</sup>。在如交通、医疗、卫生、安监、金融、社保、科技、农业、地理、环境等社会公众和市场主体关注度较高的领域, 政府数据应当优先向社会开放<sup>[19]</sup>。在数据开放过程中分为无条件开放和有条件开放两部分。其中无条件开放的政府数据要遵循公开、透明、及时、准确等原则, 应当提供给社会各界共享使用; 有条件开放的政府数据要确定数据开放的边界、范围、用途等, 这些数据仅提供授权给部分用户共享使用。

数据统计表明, 以下政策中开放数据主题提及较高:《中华人民共和国政府信息公开条例》《关于印发促进国土资源大数据应用发展实施意见》《贵阳市政府数据共享开放条例》《银川市城市数据共享开放管理办法》《苏州市大数据产业发展规划》。从开放数据共享角度来说, 这些政策文件可为政策制定者提供决策依据。

### 4.2 构建政府数据安全保障体系

政府要从构建政府数据安全保障体系角度着手提升数据安全在政策文件中的权重, 政府应从制度建设、应急管理、安全监管、人才培养、知识产权等方面尽快构建完善的数据安全保障体

系,给各行业的营商活动提供安全保障环境。数据安全是开放数据的核心能力,要通过构建数据安全保障体系切实维护公共安全和社会稳定<sup>[20]</sup>:一是积极推进数据安全立法相关工作,通过政策法规明确数据安全主体责任;二是加强数据安全监管执法的力度;三是运用智能技术手段,真正确保数据信息的可信性和安全性。

数据统计表明,以下政策中数据安全主题提及较高:《中华人民共和国网络安全法》《国务院关于大力推进信息化发展和切实保障信息安全的若干意见》《大数据安全标准化白皮书》《大数据安全白皮书(2018年)》《贵阳市大数据安全管理条例》《深圳市网络与信息安全突发事件应急预案》《广东省深入实施知识产权战略推动创新驱动发展行动计划》。从保障数据安全角度来说,这些政策文件可为政策制定者提供决策依据。

## 5 结语

本文通过对数据政策中的开放数据与数据安全子系统的深入分析,在此基础上建立了复合系统协同测度模型,并对其有序度及复合系统协同度进行计算,最终为两者协同发展提供了科学化的理论支撑和数据参考<sup>[21]</sup>。在实践意义上,政府相关部门可根据该模型实现对我国各省份数据政策中开放数据与数据安全复合系统发展程度的动态监控,为政策制定者优化现有政策提供可资借鉴的客观依据。该模型具有较好的普适性,可用来对我国数据政策中两类主题及两类政策间协同发展程度进行动态监控,为政府推进数据政策有关决策及政策制定提供科学依据和参考。

### 【参考文献】

- [1] 孙瑞英,马海群.太极哲理对开放数据与数据安全政策协同的启示[J].现代情报,2018,38(5):3-10,24.
- [2] 毛子骏,郑方,黄膺旭.政策协同视阈下的政府数据开放研究[J].电子政务,2018(9):14-23.
- [3] 王德庄,姜鑫.科学数据开放政策与个人数据保护政策的政策协同研究——基于利益相关者理论视角[J].情报资料工作,2019,40(3):39-45.
- [4] 丁红发,孟秋晴,王祥,等.面向数据生命周期的政府数据开放的数据安全与隐私保护对策分析[J].情报资料工作,2019(7):151-159.
- [5] 张彬,彭书楨,金如焯,等.“大智物云”时代数据治理国家战略比较分析——数据开放、网络安全保障与个人隐私保护[J].电子政务,2019(6):100-112.
- [6] 孟庆松,韩文秀.复合系统整体协调度模型研究[J].河北师范大学学报,1999(2):38-40,48.
- [7] HAKEN H. Synergetics. Introduction and advanced topics [M]. 3rd Berlin: Springer, 2004:24-45.
- [8] 李良成,陈欣,郑石明.科技人才与科技创新协同度测度模型及应用[J].科技进步与对策,2019(10):130-137.
- [9] 闫倩,马海群.我国开放数据政策与数据安全政策的协同探究[J].图书馆理论与实践,2018(5):1-6.
- [10] 张晓青,盛小平.国外科学数据开放共享政策述评[J].图书馆论坛,2018(8):147-154.
- [11] 王宏起,徐玉莲.科技创新与科技金融协同度模型及其应用研究[J].中国软科学,2012(6):129-138.
- [12] 夏业领,何刚.中国科技创新—产业升级协同度综合测度[J].科技管理研究,2018(8):27-33.
- [13] 马海群,张涛.文献信息视阈下面向智慧服务的语料库构建研究[J].情报理论与实践,2019,42(6):124-130.
- [14] 汪涛,谢宁宁.基于内容分析法的科技创新政策协同研究[J].技术经济,2013(9):22-28.
- [15] 张涛,马海群.一种基于LDA主题模型的政策文本聚类方法研究[J].数据分析与知识发现,2018,21(9):59-65.
- [16] ROTHWELL R, ZEGVELD W. Reindustrialization and technology [M]. London: Logman Group Limited, 1985:83-104.

[17] 樊华, 陶学禹. 复合系统协调度模型及其应用 [J]. 中国矿业大学学报, 2006(4): 515-520.

[18] 张涛, 马海群. 基于政策文本计算的开放数据与数据安全政策协同研究 [J]. 情报理论与实践. <http://kns.cnki.net/kcms/detail/11.1762.G3.20191230.1422.012.html>.

[19] 马海群, 洪伟达. 我国开放政府数据政策协同的先导性研究 [J]. 图书馆建设, 2018(4): 61-68.

[20] 彭知辉. 论政府在大数据发展中的作用: 以大数据政策为视角 [J]. 广东行政学院学报, 2017(1):12-20.

[21] 李江, 刘源浩, 黄萃, 等. 用文献计量研究重塑政策文本数据分析——政策文献计量的起源、迁移与方法创新 [J]. 公共管理学报, 2015,12(2):138-144.

## Research on the Synergy Degree of Open Data and Data Security Based on China's Data Policy

MA Haiqun<sup>1</sup> ZHANG Tao<sup>2</sup> YI Yang<sup>3</sup>

( 1. Research Center of Information Resource Management, Heilongjiang University, Harbin 150080,China;

2. Information and Network Center, Heilongjiang University, Harbin 150080,China;

3.Department of Mathematics, Heilongjiang University, Harbin 150080,China )

---

**Abstract:** [ **Purpose/significance** ] For a long time, the opening and protection, sharing and security of data are the severe challenges in the process of making and implementing big data policies. It is of great significance for the government to optimize the allocation of information resources and promote the innovation of big data industry to explore the orderly and coordinated development of open data and data security in data policies.

[ **Method/process** ] This article firstly calculates the degree of order of the open data subsystem and the data security subsystem through the method of policy text calculation, and then constructs a synergy degree model of the two composite systems, and finally analyzes the interaction between open data and data security in data policy, and expounds the mechanism of their collaborative development. [ **Result/conclusion** ] Based on data analysis of the 446 data policies collected, we can see that the benign cooperative development mechanism between the open data and data security subsystems has not been formed in China's data policies, but the development trend of the orderly and collaborative degree between the open data and data security themes in the trunk and branch policies is good, which also shows that the policy adjustment meeting will make the open data and data security have a better collaborative relationship, the paper finally puts forward the countermeasures and suggestions to promote the collaborative development of the composite system.

**Keywords:** Open data; Data security; Policy synergy; Policy text computing

---

( 本文责编: 王秀玲 )