

# 政府开放数据故事化实现模式研究

王萍 周霞

(吉林大学管理学院, 长春 130022)

**摘要:** [目的/意义] 数据故事化通过叙事方式和可视化技术揭示了数据的内在逻辑关联, 支持全景式描述数据, 将数据故事化引入到政府开放数据实践中, 有利于提高用户对数据的感知、理解、探索和重用, 进而识别、挖掘出数据隐性价值。[方法/过程] 以美国的开放数据平台为例, 从主体、基础、工具、客体4个层面分析并讨论数据故事化的实现框架。[结果/结论] 依据“5W”传播理论提出了政府开放数据故事化的运行模式, 即“创作者-数据-政府开放数据平台-受众-评价”, 以及保障数据故事化运行的4种全方位协作方式: 精细化循环把控数据全流程、嵌入自动生成故事化模块、采用区块链技术智能监测数据以及支持融合多源渠道数据的能力。

**关键词:** 政府开放数据 数据故事化 数据叙事 数据价值

**分类号:** G203

**DOI:** 10.31193/SSAP.J.ISSN.2096-6695.2020.02.01

## 0 前言

全球诸多国家政府开放数据取得了突破性进展, 较多政府开放数据平台应用可视化技术增强数据的视觉感知力, 但目前政府开放数据的可视化功能还停留在数据显示层面, 未深入挖掘、揭示及标引数据上下文之间隐藏的逻辑关系。再者, 政府开放数据仍存在着缺乏关联、难以理解、用户缺乏积极性等问题, 导致政府开放数据的隐性价值难以凸显。尽管政府数据管理部门已通过开展各类创新大赛鼓励用户参与, 但收效甚微。研究表明: 有意义的故事比统计数据更容易让人理解和记忆<sup>[1]</sup>, 故事表征了内容或事件背后的逻辑及关联关系, 具有通俗性和易懂性, 能够调动用户的积极性。政府开放数据故事化有可能解决数据隐性、费解等难题, 成为激发用户灵感的重要途径。根据文献调查, 政府开放数据与数据故事化结合的相关研究还比较缺乏, 主要针对将数据故事化功能扩展到政府开放数据平台<sup>[2]</sup>, 数据的故事化可能成为政府开放数据平台未来的发展趋势。数据故事化(Data Storytelling)具备3个基本要素: 叙事、数据、视觉效果, 借助可视化技术和文本叙事方式将

**[作者简介]** 王萍 (ORCID: 0000-0002-0308-7362), 女, 博士生导师, 教授, 博士, 研究方向为科技信息服务、信息资源管理, Email:jlwang65@163.com; 周霞 (ORCID: 0000-0002-1440-4394), 女, 博士研究生, 研究方向为政府数据开放, Email:zhouxia0129@163.com。

数据融入到叙述中并匹配特定场景,揭示数据集之间的逻辑关联,赋予数据发声、解释、说服和引导作用,便于用户洞察数据的隐形关系,从而发现数据潜在价值<sup>[3]</sup>。随着用户需求的动态演变,政府引入数据故事化的理念、方法、技术可以满足用户体验和提升平台开放数据的服务能力。结合数据新闻和数据科学中数据故事化的内容,本文运用案例、网站调查和移植等方法探索政府开放数据故事化的运行模式及实现方式,以期丰富政府开放数据故事化的理论。

## 1 研究现状

学术界关于数据故事化的术语称谓并不统一,相关解读涵盖数据驱动故事、数据叙事、用数据讲故事等。国外数据故事化的研究主要集中在理论分析、方法探析、以及实践应用等方面。(1)理论分析:Tim Davies等<sup>[4]</sup>认为,访问原始数据并不能增加重用数据的概率,然而创建数据的上下文有利于用户加强对数据源及如何使用数据的理解,进而增加数据的使用频率;Marjanovic等<sup>[5]</sup>运用边界对象的理论基础,认为数据故事化能更好的共享并迁移潜在的知识。(2)方法探析:Matthew等<sup>[6]</sup>利用时间轴来作为数据故事叙事的手法,通过时间表的设计、空间辅助与叙事点的相互匹配,有利于讲故事的人将数据形成一个跨越不同叙事点的连贯故事,进一步促使用户理解数据并传达数据背后的价值;Javier等<sup>[7]</sup>通过数据叙事游戏的设计与评估,提出了数据故事化的定性(领域专家)与定量(机器学习)分析;Adegboyega等<sup>[3]</sup>总结出数据故事化需要关联规则挖掘、数据融合和集成、机器学习、遗传算法、预测建模、情感分析、空间分析、时间序列分析等方法,辅之以R、Python、Tableau、Google融合等软件工具,以实现数据故事化。(3)实践应用:Niall等<sup>[2]</sup>倡议将数据驱动的故事化功能扩展到政府开放数据平台,以更好的衔接数据可视化与数据故事化的关系,帮助用户更好的“向下钻取”数据并获取丰富的相关元数据信息,如有关出处、异常值、突出值等。

国内数据故事主要集中在数据科学以及数据新闻领域。朝乐门等<sup>[8]</sup>阐述了数据故事化的要素、内涵、结构、方法等,为数据故事化的进一步发展奠定了基础。许向东<sup>[9]</sup>提出了新闻叙事的3种主要方法:线型、组合、交互型叙事,为相关学科探索数据故事化的实践提供了借鉴作用。李岩等<sup>[10]</sup>以《数据新闻手册》为分析材料,通过语料库及文本分析方法,探讨了数据新闻中好故事的理念:即数据驱动的可视化故事、数据故事应该具备把握真实及揭露真相等特征。

综上,国内外数据故事化的研究尚处于起步阶段。从宏观上看,数据故事化的研究较少,较多集中于数据科学、数据新闻等领域,管理学、教育学也略有涉及<sup>[11][12]</sup>。从微观上看,国内外学者普遍认为,Web技术开发与发布、数据分析与数据可视化是数据故事化的核心技术,数据故事化是解答数据之间的因果、关联、聚类逻辑联结关系,但在数据故事化展示过程中出现了重技术轻数据关系的现象。如在数据新闻领域,存在数据故事可视化炫技趋势,而忽视了数据叙事的文本情境,达不到揭露数据内在逻辑和潜在价值的目的。因此,政府将数据故事化引入开放数据中,需保障故事叙事和技术方法的平衡,从而实现政府开放数据得以优化的初衷。

## 2 美国开放数据故事化个案剖析

单案例因与研究对象具有一定的相似性,可以揭示特定现象的复杂机理<sup>[13]</sup>。为了探索数据故

事化如何融入到政府开放数据中, 本文选择圣克拉拉县开放数据平台 (County of Santa Clara Open Data Portal, 简称 CSCODP)<sup>[14]</sup> 作为案例研究, 主要原因有: (1) 根据布鲁金斯学会数据, 圣克拉拉是硅谷的故乡, 人均 GDP 排名世界第三 (仅次于瑞士的苏黎世和挪威的奥斯陆), 财富的集中使其成为美国西海岸最富裕的县, 也是美国最富裕的地方之一, 在数据开放平台上已有 5 个故事, 可作为研究样本; (2) CSCODP 的故事符合数据故事化的结构要素, 可作为深入研究的案例。

## 2.1 CSCODP 概况

CSCODP 于 2018 年 2 月由 Socrata 开发<sup>[15]</sup>, 有 8 种数据类型: 商业、教育、环境、政府、健康、住宅开发、公共安全、交通, 并设置了日历、图表、数据透视、数据集、外部数据集等不同查阅方式 (View Types)。不同数据类型可用的查阅方式见表 1。

表 1 CSCODP 数据概况

数据 类型	查阅 方式	日历	图表	数据 透视	数据集	外部 数据集	文本 文件	视图	形式	地图	故事	查阅数 据总量
商业		×	×	×	×	×	√	×	×	×	×	1
教育		×	×	×	×	×	×	×	×	√	×	3
环境		×	×	×	×	√	×	×	×	√	×	5
政府		×	√	×	√	√	×	×	×	√	√	46
健康		×	√	×	√	√	×	×	×	√	√	14
住宅开发		×	×	×	×	×	×	×	×	×	×	0
公共安全		×	√	×	√	×	×	×	×	×	×	4
交通		×	×	×	×	×	×	×	×	√	×	4

如表 1 所示, CSCODP 中设置了故事查阅方式, 但仅有政府及健康类数据可用。此外, 这两类数据还同时支持图表 (Charts)、数据集 (Datasets)、外部数据集 (External Datasets) 等查阅方式, 表明数据故事化可以和其他数据表征方式同时使用。

截至 2020 年 2 月 13 日, CSCODP 中符合数据故事化要素特征的故事有 5 个, 见表 2。

表 2 CSCODP 故事调查表

故事名称	故事来源	更新时间	访问次数	故事化展现元素
儿童预算	政府	2019/12/19	570	①附原始数据集; ②创作者叙述; ③提供与社交网络共享接口; ④视觉增强, 附各类图表
与当地企业合作, 确保 CSC 的食品安全	健康	2019/10/23	2133	①创作者驱动叙述; ②附源数据链接; ③视觉增强, 附各类图表
健康美甲	未分类	2019/08/30	1244	①创作者驱动叙述; ②附网站链接; ③解释; ④视觉增强, 附各类图表
确保 CSC 人体艺术的安全措施	未分类	2019/12/14	1056	①创作者驱动叙述; ②附统计报告链接; ③视觉增强, 附各类图表; ④源数据链接
CSC: 以服务为导向的劳动力	未分类	2019/09/26	593	①创作者驱动叙述; ②视觉增强, 附各类图表

如表2所示,这5个故事分别来源于政府、健康、未分类。在此,归纳了5个故事的异同点:(1)共同性,5个故事均有创作者叙述、视觉增强、附各类图表,如儿童预算(Children's Budget)<sup>[16]</sup>,增加了各种颜色与样式的图表,其中所附的源数据表格(占60%)能够引导受众获取源数据,并再次创造数据价值;(2)差异性,5个故事的创作并没有统一的标准,如“确保CSC人体艺术的安全措施(Ensuring Safe Practices for Body Art in Santa Clara County Through Permitting and Inspections)”创作较丰富<sup>[17]</sup>,在故事中附统计报告链接。简言之,CSCODP融入了全新的理念,虽然平台的功能设计略微单一,但与其他查阅方式相比,更新颖,具有趣味性、画面感、联结力,可以降低用户感知负荷,在吸引受众方面具有一定的优势。

## 2.2 分析与讨论

根据朝乐门提出的数据故事化生产过程(数据-故事建模-故事叙述)<sup>[18]</sup>和Bradshaw提出的双金字塔结构(其中倒金字塔自上而下的流程为“编辑-清理-情境与综合”,正金字塔自上而下的流程为“视觉化-叙事-社交化-人性化-个性化-应用化”)<sup>[19]</sup>,再结合案例数据故事制作的过程,本文从主体层、基础层、工具层和客体层4个维度构建了数据故事化案例分析框架,见图1所示。主体层依据数据新闻<sup>[20]</sup>、数据故事化创作者(故事作者、故事叙述者)建立,他们都对数据故事化具有主导作用,一般为政府公职人员或社会公众;基础层是数据故事化的基本保障,包括数据以及呈现数据故事的平台,主体层需借助基础层创作和传播数据故事;工具层是创作者将数据进行故事化时选择使用的工具,而选择何种工具与故事呈现的效果密切相关;客体层包括受众如何接收数据故事,以及如何应用等要素,例如个性化定制。

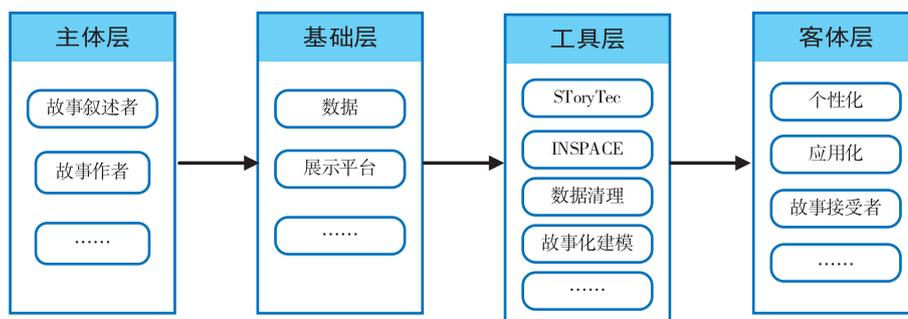


图1 数据故事化案例分析框架

### 2.2.1 案例分析

(1)主体层。由于本文的故事作者以及故事叙述者皆为政府和其委托的商业公司,故本文将故事作者以及故事叙述者统称为创作者。CSCODP中的5个故事,创作者主要为政府或委托相关商业机构(CSC开放数据由Socrata托管)。相较于传统的故事叙事方法“发展-高潮-结束”模式,创作者揭示了数据集的背景,为数据展示提供上下文,突出数据间的逻辑关系。

(2)基础层。如表1所示,CSCODP的数据内容丰富,包含了商业、教育、环境、政府、健康、住宅开发、公共安全、交通等数据集。在CSCODP上,数据集开放总量仅24个(检索时间为2018年2月至2020年3月),但提供了API应用接口,数据集维护服务较好,如平台对数据

的搜集、保密范围、技术分析等均有相关隐私申明。从平台看, CSCODP 的浏览界面简洁、可通过首页直接进入故事板块, 为故事展示提供了快捷通道, 但平台呈现内容相对简单。

(3) 工具层。如表 2 所示, CSCODP 上的故事呈现了各类图表、原始数据及报告链接, 注册用户可以选择合适的工具自定义可视化的呈现, 在 CSC 开放数据平台链接到 Socrata 的开发者中心, 为开发者提供了开发工具, 如开放源代码。因此, CSCODP 实现数据故事化满足了工具条件, 以贯彻故事比电子表格更具影响力的思想<sup>[21]</sup>。

(4) 客体层。主要是指故事受众, 分为注册和非注册受众。注册的受众能进行讨论、自定义可视化等功能, 满足一定的个性化需求。非注册受众可直接在平台上点击浏览有关数据集或者故事。

### 2.2.2 案例讨论

将优质的数据故事化模式移植于政府开放数据, 是一种有益的尝试, 然而也不可避免的存在诸多障碍, 对于案例的深度讨论, 有利于促进政府开放数据故事化的实现<sup>[22]</sup>。

(1) 生成受众画像以洞悉用户需求。创作者与受众是数据故事化最主要的参与者。数据故事化最常见的类型就是创作者主导及受众驱动的模式, 创作者主导模式是指在故事化描述的过程中, 不与受众直接交互, 仅仅是为了展示而讲故事。CSCODP 的故事多为创作者主导模式, 导致数据故事类似于说明书, 不利于平台的可持续发展。在政府开放数据运行模式中, 受众对数据故事进行接收, 并做出反馈, 可以表示为“创作者-受众”的传播形式。在传播过程中, 数据故事依托平台传递给受众, 受众可以是单个用户也可以是多用户。在 CSCODP 上, 允许用户将数据故事中的数据分享到 Facebook、Twitter 等社交媒体, 但不允许编辑故事, 受众之间可以讨论数据故事。

(2) 集成故事要素以巩固叙事基础。数据故事化主要有数据、视觉效果、叙事 3 个要素<sup>[9]</sup>。①数据, 是数据故事化最重要的元素之一, 不同领域数据来源较为不同, 如数据新闻领域, 数据主要来源于新闻工作者搜集的政府开放数据、公共数据等。②视觉效果, 为了确保受众对数据的理解与感受, 目前政府开放数据故事化采用可视化技术, 增强画面感, 降低感知负载, 如色彩、大小、饱和度以及信息图的匹配等变化<sup>[23]</sup>。③叙事, 是以文字的方式揭示数据背后故事, 使数据之间形成连贯的叙述过程。随着数据故事的逐渐发展, Edward 等<sup>[24]</sup>提出 7 种可视化的叙事形式: 杂志样式、注释图表、分区海报、流程图、漫画、幻灯片和视频, 以此实现数据故事化中创作者与受众的平衡。在 CSCODP 中, 数据来源较为丰富, 根据受众需求呈现可视化元素为数据故事化提供了良好的基础。

(3) 融合技术方法以全面解读数据。叙事包括空间、时间、视角等内容, 是一种传播事件的方法<sup>[25]</sup>。传统的故事化遵循“发展-高潮-结束”的模式来建构全文, 而数据故事化表现为非线性的叙事方法<sup>[26]</sup>。笔者认为政府开放数据故事化与数据新闻保持着同样的叙事模式。许向东<sup>[9]</sup>总结出目前数据新闻叙事模式主要有三种: ①线性叙述, 可以追踪反映客观事件的发展变化, 呈现单向性、直线型的发展特征, 较多表现为按时间发展进行叙述; ②组合性叙事, 主要为组合各种相对完善的“故事模块”, 使受众可以选择性进行阅读; ③交互性叙事, 交互性创作者通过场景、路径、体验等设计, 以满足受众的个性化体验。在数据新闻中, 优质的数据故事通常混合了线性、组合性、交互性等叙事模式, 表现形式多样, 如新华网数据新闻便呈现出声音、视频、漫画、图表、叙述等一体化的特征<sup>[26]</sup>。CSC 开放数据故事以叙事文字加图表, 以碎片化组

合叙事方法呈现数据，方法略为单一。较为成功的数据故事化通常结合多种叙事手段，如线性、组合、交互叙事等，以凸显数据故事的丰富性。

通过 CSCODP 案例的分析讨论，CSCODP 数据故事与数据新闻相比，存在着巨大的差距。数据新闻是数据故事最成功的领域，也称为数据叙事、用数据讲故事、数据驱动故事化等，其核心是新闻叙述以及如何用数据讲好故事，通过讲故事的方式报道新闻，让事实信息更具易读性和有趣性。“人民网图解新闻”、“新华网数据新闻”、“财新网数字说”等均是数据新闻制作的典范。在数据新闻实践中，优质的数据故事以及数据叙事的可用性越来越强，因而可以移植优质故事为政府开放数据使用。政府开放数据故事化侧重于受众对数据故事的接纳程度，探讨叙事故事（传统）与数据叙事的对应关系，有利于挖掘数据叙事的优势，进而形成政府开放数据故事化特有的运营流程。唐雪杰<sup>[25]</sup>从“who、when、where、how、叙事点”角度分析了数据叙事的优势，叙事是某人讲述含有时间、空间、地点等故事的一种传播活动，也是叙事者与读者共同探讨故事规律的活动，而数据叙事继承了叙事的所有方法及理论，其实现流程主要针对所搜集的数据进行统计、分析、可视化的处理之后确定叙事点，因而数据叙事并不是单纯的数据论证，而是数据与故事的融合模式。见图 2。由此可见，数据叙事不受时间、空间的限制，并且具备预测未来的属性，能够把复杂的事物“简化”，对用户抓住关键信息具有重要意义。

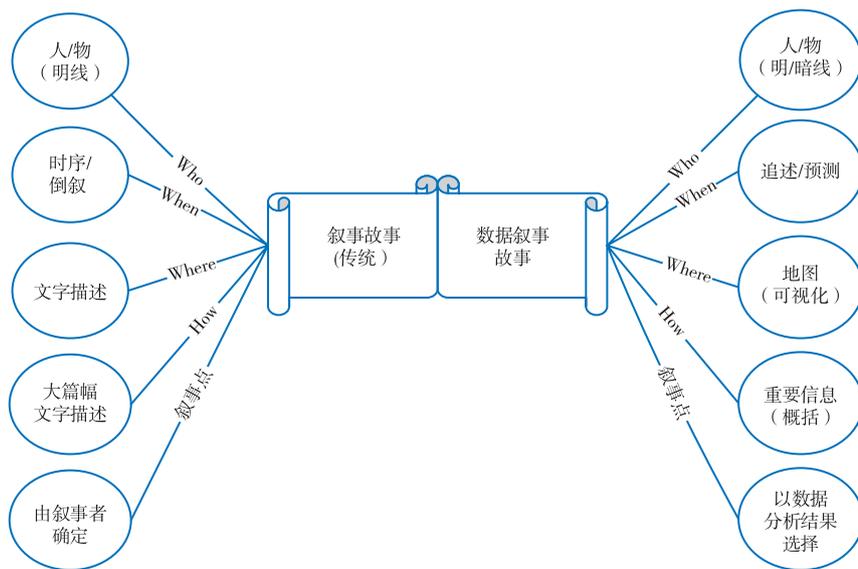


图 2 叙事故事（传统）与数据叙事故事的对应框架

目前，推进政府开放数据故事化过程中仍存在着以下问题：①缺乏理论指导。通过国内外文献调查发现，政府开放数据与数据故事化研究的结合屈指可数，缺乏成熟的理论体系。②需要转变观念。当前，对于政府开放数据而言，数据故事化方式并不会改变其本质，只是通过叙述及视觉可视化方式改变了数据的呈现形式。需要认识到，政府主导的开放数据，究其根本是为了实现数据价值，创造社会效益，如果以易于理解的数据加以呈现，有利于揭示深层的数据关系，最终

加速数据价值的实现。③技术有待加强。目前, 政府开放数据故事缺少视频、音频、互动交流、隐藏信息等功能, 还未实现智能化的生成路径。④缺乏对受众需求的及时追踪。受众的需求是政府开放数据的风向标, 满足受众的需求是推动政府开放数据故事化发展的主要因素之一。

### 3 政府开放数据故事化的运行模式

本文从宏观及微观层面分析政府开放数据故事化的模式, 确保全面的探索政府开放数据故事化的实现。政府开放数据故事化是一个复杂的系统, 不仅贯穿于数据的生命周期, 也与数据环境密不可分。Lorenz<sup>[28]</sup>提出, 数据故事化的流程为“数据-过滤-可视化-故事化”。政府开放数据依然遵循新闻叙事流程, 政府开放数据的本质决定了数据搜集及过滤程序相对简单, 其中可视化作为故事化的展示方式, 是实现故事化的要素之一。复旦大学数字与移动治理实验室发布的《2019 中国地方政府数据开放报告》显示, 在 82 个地方政府开放数据平台中, 具有丰富的图表和知识图谱等可视化展示功能的平台仅占 20%<sup>[29]</sup>。具体而言, 多数平台的可视化功能为导出图像、在线演示, 缺乏编写故事的功能。从整体上看, 我国还未建立统一的政府开放数据平台, 地域之间数据无法关联, 可视化呈现单一区域性特点, 不利于数据故事全景化扫描的实现。为了实现政府开放数据故事化运行, 本文将 Harold 提出的 5W 传播理论应用于政府开放数据的实现模式, 即谁 (Who)、说什么 (What)、通过什么渠道 (in Which channel)、对谁说 (to Whom)、取得什么效果 (with What effect), 其对应的研究内容为: 控制-内容-媒介-受众-效果<sup>[30]</sup>, 并据此提出政府开放数据故事化的运行模式, 表现为“创作者-数据-政府开放数据平台-受众-评价”, 见图 3。

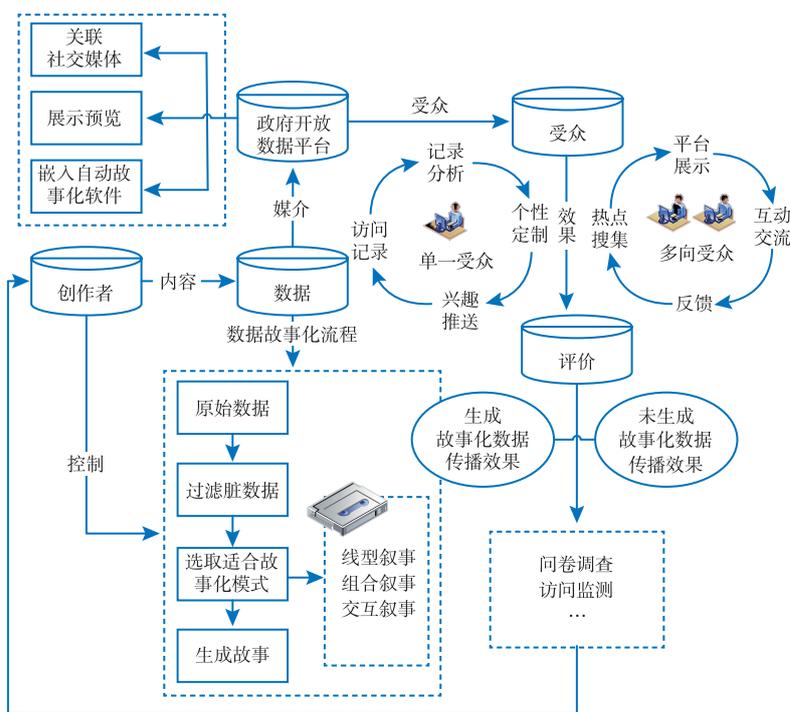


图 3 政府开放数据故事化运行模式

### 3.1 精细化循环把控数据故事化全流程

该循环模式保障了政府开放数据故事化全流程的控制。如图3所示：①政府开放数据故事化的模式驱动者即创作者主要给政府提供数据来源。由于政府开放数据要求具有客观性，而数据故事化会加入个人主观色彩，因而在创作过程中，需要对政府开放数据施加一定控制，如过滤脏数据、选取适合的故事化模式等。②政府开放数据具有丰富的基础数据，创作者应用适当的方式，以保障在搜集、整理、储存、故事化制作等环节中数据内容的可靠、完整、真实。③平台扮演了媒介的作用，是展示创作者数据故事的舞台，也是受众获取数据故事的窗口。为了保证受众群体最大化，允许将平台链接至相关的社交媒体进行展示预览，以拓宽政府开放数据故事平台的用户覆盖面。由于政府开放数据体量庞大，仅依靠人工制作数据故事无法实现真正的智能化，故而在平台中嵌入数据故事化自动生成模块。④本文将受众分为单一受众和多向受众：单一受众是个体接受者，具有个性化的特点，其流动方式主要表现为“创作者-受众”的单一循环传播，其优势在于能够快速获取个体特点，给予定制服务；多向受众是指多个个体共同获取故事，具有趋同性的特点，其流动方式为“创作者-受众-受众”的多向传播。⑤政府开放数据故事化的成效评价，对已生成故事化的数据及未生成故事化的数据均进行评价，采用问卷调查、访问监测等多种方式，为政府开放数据故事化的成效管理提供参考样本。

### 3.2 嵌入数据故事化自动生成模块

政府开放数据平台已经实现了检索、互动交流、API接口、可视化、成果展示等功能，对政府开放数据故事化的实现具有重要意义。目前，在数据新闻领域，较多使用各种软件制作数据故事，以降低人力财力成本，例如，广泛应用创作工具StoryTec、INSPACE的可视化方式创作故事。目前，我国政府开放数据平台形成了省级-副省级-地市级的分布，开放的数据集达到71092个<sup>[31]</sup>，若依靠人工制作数据故事，势必会造成效率低、数据遗漏等问题。另外，由于受众的需求总是随着社会的发展、周围环境等因素动态的变化，要准确的把握这些需求，仍然会给政府开放数据故事化的推进带来挑战。因此本文提出两种嵌入路径：①在平台中嵌入数据故事化软件，实现自动故事化生成机制，允许用户围绕数据集构建数据故事，不仅可以提高数据故事化处理的效率，也可以增加创作者与受众之间的互动；②将发现的事件插入到预定义的故事模板中，集合已有的数据集，自动生成数据故事，以推动政府开放数据故事化的可持续发展。对于政府开放数据而言，在自动化生成数据故事的过程中，通过受众的浏览、下载、分享等行为分析，以此反推需求量较大的政府开放数据，从而真正的实现数据价值。

### 3.3 采用区块链技术智能监测数据

在该实现模式中，包含两个智能点，即：提供应用数据的创作者如何规避数据质量低下、数据污染等问题，以及创作过程中如何使用正确有效的工具保障应用数据的全面科学性。对于创作者而言，在生成政府开放数据故事的过程中，其制作会掺杂个人的主观情感，特别是互动类的故事，创作者为了吸引受众，有时会通过篡改数据达到夸张的效果，因而应设立监管机构，定期监测创作者的数据故事，以保障数据故事化的过程科学透明。对故事作者而言，选择合适的创作工具利于高效的呈现数据故事，由于政府开放的数据并非全部适合于数据故事的制作，如果过分提炼，可能会导致创作者及受众受困于信息茧房。随着人工智能、区块链技术的不断发展，积极的

将其引入到政府开放数据故事化的应用中, 可以高效的管理数据。区块链具有智能合约、防篡改等特点, 不仅便于管理庞大的政府开放数据, 也能保证数据的真实性<sup>[32]</sup>。

### 3.4 融合多渠道数据实现全景式故事描述

数据融合是指基于特定的方法分析不同来源的数据, 以保证数据的全面性以及分析结论的客观性<sup>[33]</sup>。对政府开放数据故事进行全景式描述, 需要多层面的数据, 包括公共数据、商业数据等。例如, 制作数据新闻时, 除搜集政府开放数据外, 也搜集与制作故事有关的数据集, 以保证数据故事化呈现的真实性。政府开放数据的本质决定了其仅开放与公共事务密切相关的数据<sup>[34]</sup>, 因此需要关联更多相关数据, 以实现数据的共享与重用, 为数据故事化奠定基础。此外, 为保持政府开放数据的原本属性, 应将关联的数据加以区分, 形成“聚合-分解”结构, 如链接原始政府开放数据表格。

## 4 结语

政府开放数据对社会经济具有积极的影响, 对政府开放数据的研究一直是各领域的热点。相比于纯粹的数据而言, 以故事的形式呈现更容易令受众产生感性认知, 从而发现数据的多重价值。本文通过案例分析, 发现政府开放数据故事化仍处于探索阶段, 并尝试根据数据新闻、数据科学、5W 理论构建政府开放数据故事化的运行模式, 呼吁我国政府开放数据积极的引入故事化模式。然而, 在将故事化方法应用到政府开放数据的实践中, 仍需结合计算机、数据科学、新闻学等多学科的理论及方法, 以保证故事化的数据不受主观情感影响, 防止数据污染。另外, 由于故事化对于政府开放数据是新事物, 因而本研究也存在着一定的局限性: ①选取的案例样本有限, 可能对政府开放数据故事化剖析不够深入; ②政府开放数据故事化的运行可能会受到更多因素的影响, 如数据是否适合故事化处理, 如何评估数据故事化成效等问题。因此, 后续仍需从多维度考虑影响政府开放数据故事化的实现因素, 以增强研究的科学性、全面性、可行性。

### 【参考文献】

- [1] MALITA L, MARTIN C. Digital storytelling as web passport to success in the 21st century [J]. Procedia-Social and Behavioral Sciences, 2010, 2(2): 3060-3064.
- [2] BROLCHÁIN N Ó, PORWOL L, OJO A, et al. Extending open data platforms with storytelling features [C]. Proceedings of the 18th Annual International Conference on Digital Government Research. 2017: 48-53.
- [3] OJO A, HERAVI B. Patterns in award winning data storytelling: Story types, enabling tools and competences [J]. Digital Journalism. 2018, 6: 693-718.
- [4] DAVIES T, FRANK M. There's no such thing as raw data' exploring the socio-technical life of a government dataset [C]. Proceedings of the 5th annual ACM web science conference. 2013: 75-78.
- [5] MARJANOVIC O. Empowering business users to explore visual data through boundary objects and storytelling [C]. 2016 49th Hawaii International Conference on System Sciences (HICSS). IEEE, 2016: 5032-5041.
- [6] BREHER M, LEE B, BACH B, et al. Timelines revisited: A design space and considerations for expressive storytelling

- [ J ].IEEE Transactions on Visualization and Computer Graphics,2016,23(9):2151-2164.
- [ 7 ] GOMEZ J,JACCHERI L,MARAGOUDAKIS M,et al.Digital storytelling for good with Tappetina game [ J ]. Entertainment Computing,2019,30:100297.
- [ 8 ] 朝乐门,张晨.数据故事化:从数据感知到数据认知 [ J ].中国图书馆学报,2019,45(5):61-78.
- [ 9 ] 许向东.转向、解构与重构:数据新闻可视化叙事研究 [ J ].国际新闻界,2019,41(11):142-155.
- [ 10 ] 李岩,李赛可.数据新闻:“讲一个好故事”?——数据新闻对传统新闻的继承与变革 [ J ].浙江大学学报(人文社会科学版),2015,45(6):106-128.
- [ 11 ] ECHEVERRIA V,MARTINEZ-MALDONADO R,SHUM S B.Towards data storytelling to support teaching and learning [ C ].Proceedings of the 29th Australian Conference on Computer-Human Interaction.2017:347-351.
- [ 12 ] STRIDE A,FITZGERALD H F,ALLISON W.A narrative approach:The possibilities for sport management [ J ]. Sport Management Review,2017,20(1):33-42.
- [ 13 ] EISENHARDT K M.Building theories from case study research [ J ].Academy of Management Review,1989,14(4):532-550.
- [ 14 ] County of Santa Clara open data portal [ EB/OL ]. [ 2020-02-12 ].<https://data.sccgov.org/>.
- [ 15 ] Open and transparent county government: A centralized repository for public data [ EB/OL ]. [ 2020-02-23 ].<https://data.sccgov.org/stories/s/Open-Data-in-Santa-Clara-County-04-27-2018/eujc-kz27/>.
- [ 16 ] Children's budget 2019: An investment in Santa Clara county children's well-being [ EB/OL ]. [ 2020-02-13 ].<https://data.sccgov.org/stories/s/h3ir-52ge>.
- [ 17 ] Ensuring safe body art through permitting and inspections in Santa Clara county [ EB/OL ]. [ 2020-02-13 ].<https://data.sccgov.org/stories/s/n93r-endz>.
- [ 18 ] 朝乐门.数据科学 [ M ].北京:清华大学出版社,2016:98.
- [ 19 ] BRADSHAW Paul. The inverted pyramid of data journalism[EB/OL].[2020-05-09].<https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>
- [ 20 ] 刘英华著.数据新闻实战 [ M ].北京:电子工业出版社,2016:9.
- [ 21 ] Connecting data to residents through data storytelling [ EB/OL ]. [ 2020-03-06 ]<https://www2.deloitte.com/us/en/insights/industry/public-sector/chief-data-officer-government-playbook/open-data-success-stories.html>.
- [ 22 ] Strategic storytelling with open data. [ EB/OL ]. [ 2020-03-06 ]<https://govex.jhu.edu/wiki/strategic-storytelling-with-data-focusing-on-the-things-that-matter-to-residents/>.
- [ 23 ] ECHEVERRIA V,MARTINEZ-MALDONADO R,GRANDA R,et al.Driving data storytelling from learning design [ C ].Proceedings of the 8th International Conference on Learning Analytics and Knowledge.2018:131-140.
- [ 24 ] SEGEL Edward,HEER Jeffrey.Narrative visualization: Telling stories with data [ J ].IEEE Transactions on Visualization and Computer Graphics,2010,16(6):1139-1148.
- [ 25 ] 唐雪杰.数据新闻的叙事研究 [ D ].保定:河北大学,2019.
- [ 26 ] 张馨月.新华网数据新闻故事化叙事研究 [ D ].海口:海南师范大学,2019.
- [ 27 ] 数据新闻:京东告诉你大家度假的不同“姿势” [ EB/OL ]. [ 2020-02-14 ].[http://www.jjckb.cn/2019-10/09/c\\_138457618.htm](http://www.jjckb.cn/2019-10/09/c_138457618.htm).
- [ 28 ] LORENZ M.Data driven journalism: What is there to learn [ C ].IJ-7 Innovation Journalism Conference.2010: 7-9.
- [ 29 ] 复旦大学数字与移动治理实验室.中国地方政府数据开放报告 2019 上半年 [ EB/OL ]. [ 2020-02-22 ].<http://www.dmg.fudan.edu.cn/?p=6724>.
- [ 30 ] Lasswell H D. The structure and function of communication in society[J]. The communication of ideas, 1948, 37(1): 136-139.

[31] 复旦大学数字与移动治理实验室. 中国地方政府数据开放报告 2019 下半年 [EB/OL]. [2020-02-26]. <http://www.dmg.fudan.edu.cn/?p=7377>.

[32] 刘芷欣. 基于区块链技术的全流程学术评价研究 [J]. 文献与数据学报, 2019, 1(3): 38-44.

[33] 许海云, 董坤, 隗玲, 等. 科学计量中多源数据融合方法研究述评 [J]. 情报学报, 2018, 37(3): 318-328.

[34] 缪瑞生, 马海群. 我国政府开放数据现状问题及对策选择 [J]. 图书馆理论与实践, 2017(9): 16-20.

## Study on the Realization Model of Government Open Data Storytelling

WANG Ping ZHOU Xia

(School of Management, Jilin University, Changchun 130022, China)

---

**Abstract:** [ **Purpose/significance** ] Data storytelling reveals the internal logical relationship of data through narrative mode and visualization technology, supports panoramic description of data, and introduces data storytelling into the government's open data practice, which is conducive to improving users' perception, understanding, exploration and reuse of data, thus identifying and mining the hidden value of data. [ **Method/process** ] Taking the open data platform of the United States as an example, this paper analyzes and discusses the realization framework of data story from four aspects: subject, foundation, tool and object. [ **Results/conclusion** ] According to the "5W" communication theory, this paper puts forward the operation mode of government open data storytelling, which is "Creator-Data-Government Open Data Platform-Audience-Evaluation", and four comprehensive collaboration methods to ensure data storytelling operation: refined cyclic control of the entire process of data, embedding automatic generation of storytelling modules, intelligent monitoring of data using blockchain technology, and the ability to support the fusion of multi-source channel data.

**Keywords:** Government open data; Data storytelling; Data narrative; Data value

---

( 本文责编: 王秀玲 )