引文格式：党琦，张寒梅，陈重远，等. 大数据时代《咸宁城市地图集》的编制特点[J]. 地理空间信息, 2023, 21(4): 111-114. doi:10. 3969/j.issn. 1672-4623.2023.04.028

大数据时代《咸宁城市地图集》的编制特点

党　琦[1]，张寒梅[1]，陈重远[1]，鲁　钊[1]
（1. 湖北省地图院（湖北省地图数据应用中心），湖北 武汉 430074）

摘　要：《咸宁城市地图集》是一部具有地方特色，反映咸宁的自然、经济、文化等多方面的综合性城市地图集，随着大数据时代的到来，数据量爆炸性的增多，为《咸宁城市地图集》的编制提供了新的测绘地理信息数据基础。因此，从内容设计、数据源、数据处理、表达设计四个方面对《咸宁城市地图集》的编制特点进行阐述，同时，为使图集的编制更加高效、美观，从大数据挖掘技术和地图可视化两方面出发，对大数据时代《咸宁城市地图集》的编制进行思考。

关键词：咸宁城市地图集；大数据时代；编制特点；思考
中图分类号：P285　　**文献标志码：**B　　**文章编号：**1672-4623（2023）04-0111-04

Compilation Characteristics of *Xianning Urban Atlas* in the Big Data Era

DANG Qi[1], ZHANG Hanmei[1], CHEN Zhongyuan[1], LU Zhao[1]
(1. Map Institute of Hubei Province (Hubei Map Data Application Center), Wuhan 430074, China)

Abstract: *Xianning Urban Atlas* is a comprehensive atlas reflecting the nature, economy, culture and other aspects of Xianning. With the advent of big data era and the explosive increase of data volume, it provides a new basis of surveying and mapping geographical information data for the compilation of *Xianning Urban Atlas*. In this paper, we expounded the compilation characteristics of *Xianning Urban Atlas* from content design, data source, data processing and expression design four aspects. In order to make the compilation of atlas more efficient and beautiful, we thought about the compilation of *Xianning Urban Atlas* in the big data era from the two aspects of big data mining technology and map visualization.

Key words: *Xianning Urban Atlas*, big data era, compilation characteristics, thought about

咸宁位于湖北省东南部，是武汉、长沙、南昌三大省会城市经济区的地理中心。《咸宁城市地图集》是“中国城市地图集系列”中一部具有地方特色的综合性城市地图集，其从区位、环境、经济、科技、文化等不同维度展示了咸宁的综合面貌，是人类认知城市的工具，也是记录城市发展的珍贵档案[1-5]。

1　大数据时代《咸宁城市地图集》的编制特点

随着测绘科技不断创新发展，地理国情监测、天地图等产生了海量无缝的测绘地理信息数据。本文从内容设计、数据源、数据处理、表达设计四个方面对大数据时代《咸宁城市地图集》的编制特点进行阐述。

1.1　内容设计

《咸宁城市地图集》主要包括序图、区域图、街区详图、索引四部分。序图结合最新的专题数据等资料编制，旨在展示咸宁市行政区划、地理环境、资源环境、社会经济、城市风貌、历史文化底蕴等特色专题内容，因此，编制了行政区划、历史沿革、自然资源、综合经济及相关人文景观等主题共24幅专题图，每个专题由主图、附图、图表、文字、图片组成。

其中，区域图旨在展示咸宁市辖的咸安区、嘉鱼县、通城县、崇阳县、通山县、赤壁市6个县（市、区）的区域全貌及各行政区划组成单元，同时在县市地图上增加了对应的城区地图。

街区详图包括57幅大比例尺分幅详图，表示了主城区党政机关、学校医院、文化体育馆、商场超市、道路、水系、居民地等，详细反映咸宁市主城区的结构和面貌。

索引收录图集中街区详图内出现的全部地名。索引按地名汉语拼音排序，每一条索引的内容包括名称、图幅页码和图上检索网格代码。

1.2　数据源

传统地图集编制使用的数据源主要来源于测绘领域，如遥感影像数据、地面测量数据以及其他专题统计数据，往往数据量小、类型少、处理简单。

大数据时代，数据量爆炸性的增多，数据量可从数百TB到数百PB、甚至EB级的规模，如此庞大的数据规模，势必会产生庞杂无序的资料。因此《咸宁城

收稿日期：2021-12-10；修回日期：2022-03-15。
第一作者简介：党琦（1988—），工程师，从事地图制图、GIS数据库建设、GIS系统开发工作，E-mail：jomedy@163.com。

市地图集》的编制需获取合适的数据资料，应在了解源数据基本情况的基础上，建立支撑多源异构数据汇聚与采集的技术和工具，保障数据来源[5]。一般来说，图集编制的数据资料以政府等权威部门发布的数据为主，辅以其他网络资料。

《咸宁城市地图集》的数据资料主要来源于湖北省标准地图、天地图·湖北、地理国情监测、遥感卫星影像、统计年鉴以及其他专题统计数据。这些数据均为权威资料，且具有数据量大、种类繁多、数据结构类型均不一样的特点。

以序图中“综合经济”图幅为例，通过查找咸宁市近10年的统计年鉴中的经济数据，融合人口、产业等数据，进行数据分析。分析得出咸宁市各县（市、区）地区生产总值及人均地区生产总值、农林牧渔总产值及构成、形成规模的工业总产值、居民收入与支出、公共财政预算收入与支出、社会消费品零售总额、普教经费支出、进出口总额、旅游业总收入等专题数据信息。

图集中道路数据以天地图·湖北中的道路和铁路数据为主；水系数据以天地图·湖北中的数据为主，并参考地理国情监测数据中的部分水系面进行融合，调整水系面大小；植被数据以地理国情监测数据为准，根据分类码分别选取出绿地和林地；居民地将天地图·湖北中的居民地数据根据范围界提取。

1.3 数据处理

《咸宁城市地图集》的数据处理包括数据转换选取、数据清洗两大类。

1）数据转换选取。《咸宁城市地图集》数据源来源于多源数据，这些数据的投影、数据结构、制图综合、现势性均不同，需要根据技术要求进行转换、选取，保证数据资料的统一性、标准性。

《咸宁城市地图集》将最新的天地图·湖北、地理国情监测、其他专题资料等不同来源、不同尺度的大数据，按照统一的投影、数据结构进行建库处理后编制成图。其中，在处理投影时，需要将数据源的不同坐标系转换为咸宁的投影坐标系。

在处理资料的数据结构时，天地图·湖北中数据道路数据结构如表1所示，地理国情监测道路数据结构如表2所示，两者数据结构有相同的属性，但是其中国标分类码、道路编号属性名不同，地理国情监测中TYPE属性用于存储道路类型，天地图·湖北无该属性。所以，需要将两者属性名称进行统一，再融合为一套数据结构使用。

表1 天地图·湖北道路数据结构（部分属性）

属性名称	属性类型	属性说明
CLASID	Varchar	国标分类码
Name	Varchar	道路名称
ROUTENUM	Varchar	道路编号

表2 地理国情监测道路数据结构（部分属性）

属性名称	属性类型	属性说明
GB	Varchar	国标分类码
Name	Varchar	道路名称
RN	Varchar	道路编号
TYPE	Varchar	道路类型

不同的地图数据源制图综合的尺度不一致。《咸宁城市地图集》以《湖北省标准地图》中咸宁地图为基础地理底图，在此底图上针对不同的专题图幅，分别对境界、道路、水系、地名地址等要素进行不同的制图综合。

时空数据是指具数据有时间属性并随时间变化而变化的空间数据，现实世界中道路、水系、居民地等空间地理要素数据和一些专题数据时刻在发生着改变。为保障《咸宁城市地图集》的现势性，地理要素数据和专题数据需要整合各方资料，使用资料中现势性最强的部分。同时，一旦有最新的资料介入，需要立刻更新图集内容。例如，部分国道、省道道路数据进行了改道，需要对道路的走向进行更新；道路名称发生改变，则对道路的属性进行更新。

2）数据清洗。大数据时代图集的编制会收集到足够多的专题资料，但是与此同时，收集的资料会存在许多冗余无价值的和标准不一致的干扰数据，数据的真实性、准确性、有效性面临挑战，所以需要对数据进行清洗等操作，用于降低冗余，减少错误或者误差，提高数据准确性。

《咸宁城市地图集》中地名地址兴趣点以天地图·湖北中的数据为基础，但是该数据中夹杂着大量的小门面，这些小门面不能反映城市该片主城区面貌，且不稳定，容易变动，所以需要对该数据进行清洗，从中选取出当前图幅范围内POI稳定因子分数较高且能反映该片区域位置的地标类POI兴趣点。

稳定因子分数可根据式1计算，设POI重要度为 N_1，点位所位于图幅是否属于中心城区为 N_2，POI周围疏密度程度为 N_3，POI稳定因子分数的理论公式如式（1）所示。

$$S=(P_1\times N_1+P_2\times N_2)/(P_3\times N_3) \quad (1)$$

式中，P_1，P_2，P_3 为相关因子的影响系数。从式（1）可知，POI稳定分数 S 与POI重要度 N_1，成正比关

系，与是否属于中心城区 N_2，POI周围疏密程度 N_3 成反比关系。

根据式（1）可知，不同的地区需要不同的选取原则，主要的政府单位、村（居）委会、大型桥梁、隧道、大型超市、大型宾馆、中小学及高等院校、医院、大型商业体，主要的繁华商业街、公园旅游景点、大型餐饮等，这些POI是相对重要点位，在点位密集区，应选择稳定因子分数 S 较高，地图中能放置的前提下应尽量表示。但是当POI点位位于地图上稀疏的位置，其他如餐饮、便利店等非重要点位，也应上图表示。

3）表达设计。《咸宁城市地图集》封面、封底、整饰设计采用《中国城市地图集系列》的统一设计风格。序图作为专题地图，地图主体内容占用较大幅面，并安排于幅面的主要位置，其中地图专题使用定点符号法、线状符号法、质底法、等值线法、范围法、定位图表法、分级统计图法、分区统计图表法等方法表达。每幅专题图的图幅之间的色彩设计协调统一，同时突出了专题的主题内容，体现了图集的“时代感”。

如图1“赤壁之战”图幅所示（赤壁之战是中国历史上以少胜多、以弱胜强的著名战役之一，是三国时期“三大战役”中最为知名的一役）。

图幅从赤壁古战场、赤壁之战、赤壁之战古诗词，赤壁之战经典故事4个方面讲述。图幅中最重要的赤壁古战场景区示意图通过手绘地图的方式展示，同时以运动线法描述赤壁之战示意图。此外，还通过创意文字展示李白、孙元晏、杜牧的赤壁之战古诗词；以文字和手绘的形式展示赤壁之战经典故事。图幅的色彩与图集的整饰色系保持协调一致。



图1　赤壁之战

区域图、街区详图使用统一设计的地图符号、注记样式，并且增加手绘符号等多种表达手段，使图集更加生动形象。其中，如图2区域图、街区详图部分地图符号所示，这些地图符号在设计时，全面考虑了图幅比例尺、要素用途、要素分类、色彩等方面，能够真实、直观地反映要素对象，具有较高的辨识度，使读者查看地图时，能理解要素表达的意义，而无需查看对应的图例，增强了读者的可阅读性。



图2　区域图、街区详图部分地图符号

2 大数据时代《咸宁城市地图集》编制的思考

为使《咸宁城市地图集》编制更加高效，地图样式表达更加美观，本文从大数据挖掘技术和地图可视化两方面进行思考。

2.1 大数据挖掘技术分析数据资料

当数据量较小时，数据间缺乏联动性，所以数据本身不具备足够价值。而通过数据分析，可分析出数据含有深度的价值[6]。

《咸宁城市地图集》的编制，专题资料挖掘分析主要使用大数据并行计算和空间数据挖掘分析算法两种方式。对于大数据并行计算而言，Spark是目前使用最广泛的技术（Spark是一种专为大规模数据处理而设计的计算框架，其主要包括map和reduce两个阶段。map阶段，将数据切分成 N 个小文件，再将这些小文件分发到 M 个reduce节点中；reduce节点完成数据分析操作后，通过整合各节点的结果得出最终结果）。

以统计湖北省内各市、县、乡、镇的行政村（居）委会专题为例，资料使用2021年天地图·湖北最新的行政村（居）委会POI数据，通过Spark对该数据进行数据统计分析。天地图·湖北行政村（居）委会表结构如表3所示。

表3 天地图·湖北行政村（居）委会表结构（部分属性）

名称	类型	字段说明
IDCARD	Varchar	唯一ID号
NAME	Varchar	名称
PAC	Varchar	所属地区行政区划

首先在map阶段，将相同的PAC进行合并处理，其次将相同的PAC切分成若干小文件后分配到同一个reduce节点中进行处理。由于IDCARD的唯一性，所以只需计算PAC相同的IDCARD的数量就可得到某一地区行政村（居）委会的总数。天地图·湖北行政村（居）委会大数据分析操作流程如图3所示。

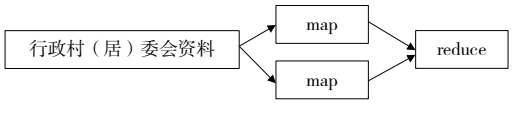


图3 天地图·湖北行政村（居）委会大数据分析操作流程

除了使用Spark大数据并行计算框架，还可以通过空间数据挖掘分析算法分析专题资料，例如文献[7]，作者提出一种基于概念格的空间关联规则挖掘优化算法。

在大数据时代，图集通过大数据挖掘技术，找出数据资料的内在联系和价值，获取专题统计信息，编制相关的专题图表等，可进一步实现地图的价值最大化。

2.2 地图可视化形式的多样性

可视化可以粗略地被定义为通过图形的表现形式，进行信息传递、表达过程。而地图可视化是将不同的地图要素以不同的形式展现的过程，可视化对于《咸宁城市地图集》的编制尤其重要。除了传统的定点符号法、分级统计图法、分区统计图表法等专题图可视化方法外，在大数据时代，还可使用Echarts、Smartbi、Tableau等可视化技术。这些可视化技术除提供常规的柱状图、饼图、K线图，还引入流数据、字符云，形成流地图等其他地图可视化展现方式，增强图面效果。同时可结合大数据分析自动生成图表，实现实时地图的展示[8]。

3 结 语

《咸宁城市地图集》是一本具有科学性的综合性城市地图集，本文从内容设计、数据源、数据处理、表达设计四个方面出发，详细介绍大数据时代《咸宁城市地图集》编制的特点，同时，从大数据挖掘技术分析和地图可视化两方面对《咸宁城市地图集》的编制进行相关的思考。

随着大数据技术的普及，未来图集编制与大数据技术会进一步相互融合。加强资料收集能力、大数据挖掘分析、数据可视化等能力，将使图集的编制更加高效、美观，为社会提供更好的地图服务。

参考文献

[1] 林媛媛，林川，何德. 浅谈大数据时代下的GIS发展[J]. 江西测绘，2013，97(3)：15-16

[2] 邓岚. 大数据时代下灾害档案数据管理探讨[J]. 经济研究导刊，2015(14)：271-272

[3] 黄洪纤.《泰州城市地图集》的设计与编制[J]. 测绘与空间地理信息，2018，41(2)：81-82

[4] 杜清运，任福，侯宛玥，等. 大数据时代综合性城市地图集设计的思考[J]. 测绘地理信息，2021，46(1)：16-20

[5] 李仕峰. 多源异构数据时空融合关键技术研究与应用[J]. 地理空间信息，2021，19(10)：19-25

[6] 范理信，党琦，鲁钊. 浅谈大数据时代下天地图·湖北的发展[J]. 地理空间信息，2018，16(4)：8-10

[7] 李宏伟，陈虎，王振宇，等. 基于概念格的空间关联规则挖掘优化[J]. 测绘科学技术学报，2013，30(3)：304-307

[8] 王瑞松. 大数据环境下时空多维数据可视化研究[D]. 杭州：浙江大学，2016