



地质大数据的特点及其在成矿规律、成矿系列研究中的应用*

王登红¹, 刘新星^{1,2}, 刘丽君^{1,2}

(1 中国地质科学院矿产资源研究所 国土资源部成矿作用与资源评价重点实验室, 北京 100037;

2 中国地质大学(北京)地球科学与资源学院, 北京 100083)

摘要 进入 21 世纪以来, 面对这一信息大爆炸的时代(即“大数据”时代), 人们的生活、工作与思维都面临着大变革。文章基于“大数据”思维, 探讨了成矿规律及成矿系列研究中的一些问题, 以期地质大数据能够在成矿规律及成矿系列的研究中得到更好地应用。首先, 从“大数据”的特点及其研究现状入手, 结合地质矿产研究中的现实问题, 阐述了地质大数据的概念及其外延。然后, 综合“大数据”与成矿规律、成矿系列研究相关各地质专业的特点, 浅析了矿产资源领域地质大数据的 10 个特点, 其中, 除了从地质矿产的视角解释了“大数据”的大量性、高速性、多样性、价值性 4 大特点外, 还基于地质矿产专业提出了地质大数据的 6 大新特点: “物质性与非物质性”“空间性与非空间性”“时间性与非时间性”“因果性与非因果性”“主体性与非主体性”及“客体性与非客体性”, 并在“大数据”的背景下作了新的诠释。最后, 总结了地质大数据在成矿规律、成矿系列、成矿体系研究中的应用情况及注意事项, 期望地质大数据能为成矿理论和成矿预测工作提供新思路。

关键词 地质学 大数据 地质大数据 成矿规律 成矿系列 成矿体系 成矿预测

中图分类号: P612

文献标志码: A

Characteristics of Big Geodata and its application to study of minerogenetic regularity and minerogenetic series

WANG DengHong¹, LIU XinXing^{1,2} and LIU LiJun^{1,2}

(1 MLR Key Laboratory of Metallogeny and Mineral Assessment, Institute of Minerals Resources, Chinese Academy of Geological Sciences, Beijing 100037, China; 2 School of the Earth Sciences and Resources, China University of Geosciences, Beijing 100083, China)

Abstract

Since the 21st Century, confronted with the epoch of information explosion (i. e., the epoch of “Big Data”), people are facing great changes in living, working and thinking. In this paper, some problems concerning the study of minerogenetic regularity and minerogenetic series of mineral deposits are discussed in the big data way of thinking in the hope that Big Geodata can be well applied to the study of minerogenetic regularity and minerogenetic series. First of all, this paper starts with the concept and research status of Big Data and, in combination with some problems concerning geology and mineral resources, expounds the concept and extension of Big Geodata. Then, after summarizing the professional characteristics of Big Data and geology and mineral re-

* 本文由中国地质调查局“中国矿产地质与区域成矿规律综合研究(中国矿产地质志)项目(编号: 1212011220369)”、国土资源部公益性行业科研专项计划“南岭东段九龙脑矿田成矿规律与深部找矿示范项目(编号: 201411050)”和中国地质大调查项目“我国离子吸附型稀土矿战略调查及研究(编号: 1212011220804)”联合资助

第一作者简介 王登红, 男, 1967 年生, 研究员, 博士生导师, 主要从事矿产资源研究。Email: wangdenghong@sina.com

收稿日期 2015-07-20; 改回日期 2015-08-29。苏 杭编辑。

sources, 10 kinds of Big Geodata characteristics have been expanded with some specific examples of geology and mineral resources. In addition to the use of geology and mineral resources perspective to explain the four characteristics (4V + 6W) of Big Data, namely volume, velocity, variety and value, six new characteristics of Big Geodata are put forward based on professional setting of geology and mineral resources, which are “what and non-what”, “where and non-where”, “when and non-when”, “why and non-why”, “who and non-who” and “whom and non-whom”, and are reinterpreted under the background of Big Data. Finally, the application of Big Geodata to minerogenetic regularity, minerogenetic series, minerogenetic system and minerogenetic prediction as well as some other aspects that deserve attention is summarized, and then some problems existing in the application of Big Geodata are also discussed, in the hope of providing research ideas for researchers of geology and mineral resources.

Key words: geology, Big Data, Big Geodata, minerogenetic regularity, minerogenetic series, minerogenetic system, minerogenetic prediction

与单个矿床的研究不同,成矿规律、成矿系列、成矿体系、成矿预测的研究,毫无疑问都需要“大数据”。那么,什么是“大数据”?什么是地质大数据?成矿规律、成矿系列等的研究需要什么样的地质“大数据”呢?“大数据”在成矿规律等的研究中应注意哪些问题呢?

1 基本概念和研究现状

在互联网上,“大数据(Big Data)”被解释为“无法在可承受的时间范围内用常规软件工具进行捕捉、管理和处理的数据集合”。互联网信息显示,“大数据”是维克托·迈尔·舍恩伯格(Mayer-Schönberger Viktor)及肯尼斯·库克耶(Cukier Kenneth)在 2008 年 8 月中旬提出的,但该词语的出现和实际使用肯定早于此时间。如,早在 1981 年,著名未来学家阿尔文·托夫勒(Toffler Alvin)便在《The third wave》一书中,将“大数据”热情地赞颂为“第三次浪潮的华彩乐章”。《Nature》在 2008 年就推出了 Big Data 专刊,《Science》在 2011 年 2 月也推出“Dealing with Data”专刊。在中国,牛文杰等在 2001 年发表了《超大数据体泛克里金插值的研究》一文;在更早的 1982 年,00289 部队的马光中等发表了《数学地质在碌曲-平武一带作成矿远景预测的应用》一文,简要介绍了在 DJ5-121 计算机上自编程序对 53 200 km² 范围进行“大数据”成矿预测的做法。

当然,《大数据时代》(Mayer-Schönberger et al., 2013)一书更具标志性。该书强调,“大数据”不采取随机分析法(抽样调查)这样的捷径,而采用所有数据进行分析处理;“大数据”具有 4V(或 5V)特点:Volume

(大量)、Velocity(高速)、Variety(多样)、Value(价值)(真实性(Veracity))。这些论述都切中要害,也是客观的,同样需要在成矿规律、成矿预测等方面的研究中加以遵守。那么,如何在成矿规律和成矿系列的研究过程中充分利用海量的数据?如何将数据与思维结合起来?如何将数据转化为新的认识或知识,并运到地质找矿的实践中?本文将对此作一探讨,以抛砖引玉,使得当前成矿规律、成矿系列乃至成矿预测的研究更具有“大数据”的时代特色。

地质大数据是信息时代背景下大数据的理念、技术和方法在地质领域的应用与实践。地质大数据涉及到地球的各个圈层,涉及地球形成与演化的历史,涉及到地球的物质组成及其变化,涉及矿产资源的形成、勘查与开发利用,涉及人类环境的破坏与修复等。在当前地质工作中,各种复杂类型数据的采集、挖掘、处理、分析与应用都与信息社会的“大数据”不谋而合,或者说,地质大数据是“大数据”的重要组成部分;而且,地质工作中实测的“实体”数据的比重可能比网络世界中通过各种各样模型计算出来的数据比重更大。这是与地质工作注重野外实践、注重第一手资料的采集分不开的。应用“大数据”理念来研究成矿系列和成矿规律,实际上就是在充分利用与“矿”有关的各种数据来厘定矿床的成矿系列,总结成矿规律并以各种适当的方式表达出来(包括声音、图像)。这种“全息”式的研究有别于传统的“抽样调查”,故也可称为“大数据成矿规律”、“大数据成矿系列”。

尽管“大数据”不是新名词,在 20 世纪 80 年代初就有人在地球科学领域的研究中提到“大数据”,但“大数据”对于当今人类生活方方面面的冲击如此

之大,显然是在2012年之后(李国杰等,2012)。2012年3月29日,美国总统奥巴马发布了《大数据研发倡议》(Big Data Research and Development Initiative),开启了美国“大数据”的全面研发,由美国联邦政府的6个部门直接承担,其中就有美国地质调查局(USGS)。中国地质学界在这方面的研究还甚少。以“大数据”为检索词,在CNKI(中国知网)上搜索,结果发现,截止到2015年7月8日16点16分,在8079种中文期刊中,在43 878 879篇文献中,直接在篇名中使用“大数据”的有9625篇,其中,2015年(尚不到1年)就占2596篇;2014年4469篇;2012年358篇;2004年13篇;2002年10篇;1999~1990的十年间只有18篇;1980~1989年仅仅5篇(图1)。可见,2010年以后才可以称为“大数据”时代。当然,在这之前并非没有“大数据”。大数据也是一个相对的说法,相对于50年以前,现今当然算“大数据”时代,但相对于50年之后呢?!

在上述所有文章的标题中同时出现“大数据”和“地质”的文章只有17篇;而同时出现“大数据”与“矿”的文章仅仅1篇。这说明,尽管“大数据”问题一开始也出现在地质科学领域(比如石油勘探),但当前的地学研究显然已经落后于其他学科!这与地质数据的海量积累、地质找矿难度的日益加大、矿产资源综合利用的要求更高很不相称。另外,地质数据往往是原始数据(诸如通过地球物理、地球化学、地质钻探等获取的数据都是原始数据,不是“加工”之后的数据)。如何对这些原始数据进行加工处理,

离不开现代“大数据”处理的新技术和新理念,而这一点恰恰又是地质科学领域所不足的。比如,在成矿年代学方面,40年以前获得一个K-Ar法数据就可能对理解成矿时代产生影响,但现今利用Ar-Ar、K-Ar、Rb-Sr、Re-Os、U-Pb等各种方法获得的一大堆数据,可能也只是为了查明某个矿床的成矿时代,深度挖掘和关联研究明显不够。

从专业数据属性来看,地学各个专业都在建立和规范管理各专业子领域的的数据资源,争取实现共享,如全国岩石学专业委员会的“火山岩数据库”,但矿床专业委员会尚未建立开放的“矿床数据库”。从矿产资源领域属性看,以矿产资源为核心(涵盖黑色金属、能源、有色金属、贵金属、稀有、稀土、稀散金属、宝玉石、非金属等各个行业),逐步拓展到相关上、下游产业(基础地质调查、矿山开采开发、矿政管理、冶金加工、环境保护、矿产品贸易等等)。从空间属性看,以中国地质调查局为主体建立全国的“大数据”,6大区、各个省级甚至地、县、跨行政区域的“成矿区带”都可以自成体系地构建“大数据”,从而为成矿区划、为矿业区域发展提供支撑。从数据来源的渠道看,既有政府部门的统计数据,也有来自于矿山企业、行业协会的各种主体经营性基本信息以及来自于媒体的投资信息、股东信息、专利信息、进出口信息、招聘信息等。但是,目前各种“数据库”都是铁将军把门、甚至“买票”也难以进入,尚未达到开放式应用的程度,也尚未同时达到开放公共数据的8条基本原则(数据必须是完整的、原始的、及时的、可读

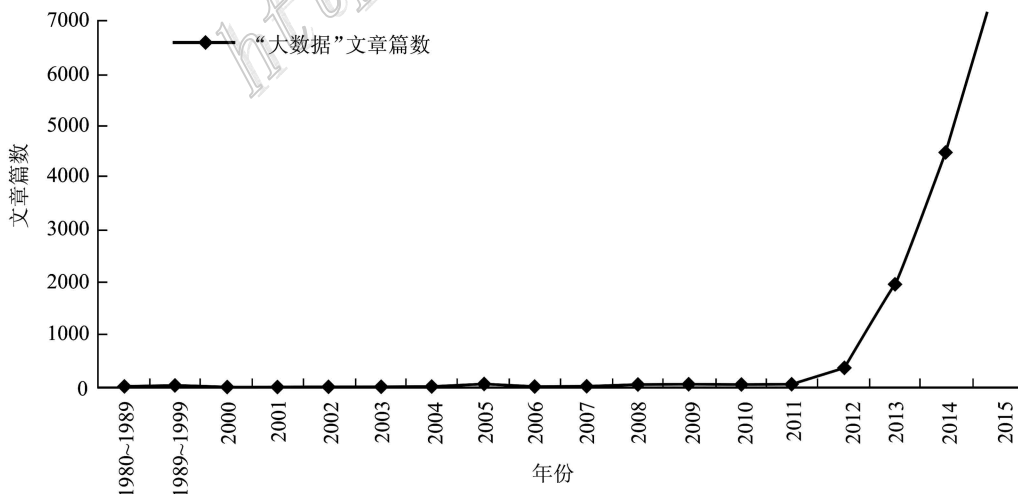


图1 近年来以“大数据”为题发表的中文论文的变化情况

Fig. 1 The changes in the number of published Chinese papers on the subject of 'Big Data'

取、机器可处理的、不需要许可的、获取必须是无歧视的、格式必须是通用而非专有的)(胡雄伟等, 2013)。

2 矿产地质领域地质大数据的几个特点

毫无疑问,矿产地质现今也进入“大数据”时代,甚至比其他领域更早。在地质尤其是矿产地质领域,“大数据”除了具有 4V 四大特点(Volume、Velocity、Variety、Value)之外,还需要强调具专业背景的 6W 特点(What、Where、When、Why、Who 和 Whom)。这 6W 是成矿规律研究的基本方法,也体现在矿产地质大数据上。

2.1 地质大数据的大量性

目前,到底有多少地质数据,并无由权威部门发布的“统一说法”,但“无法在可承受的时间范围内用常规软件工具进行捕捉、管理和处理”这一点,恐怕是不言而喻的。显然,地质数据也是一个数据集合,包括地质、矿产、遥感、物探、化探等各个专业,而且还是相互关联,融汇贯通的。仅就矿产地的数量来说,目前全国至少有 7 万处(王登红等, 2014a),在一些官方文件和科普读物中指出“已发现矿床、矿点 20 余万个……”(陈毓川等, 1999)。这 20 万处矿产地的信息是海量的,一般的笔记本电脑和台式机也难以“在可承受的时间范围内”进行日常处理。目前,一张 Excel 表格也无法容纳 7 万处矿产地的全部信息,更何况 20 万处。要对这 20 万处矿产地进行分类、排序就难以方便地完成,因此也需要借助于“大数据”的理念和技术。

目前,中国地质调查局系统已建成的数据库包括区域地质数据库(涵盖全国 1:250 万、1:100 万、1:50 万、1:25 万、1:20 万区域地质图、全国 1:20 万自然重砂、全国同位素地质测年、全国岩石地层单位等数据库)、基础地质数据库(涵盖全国岩石物性、全国地质工作程度等数据库)、矿产资源数据库(涵盖全国矿产地、全国矿产资源利用调查矿区资源储量核查成果、全国大中型矿山、矿产资源远景调查、全国主要固体矿产大中型矿山资源潜力调查、全球地质矿产等数据库)、油气能源数据库(涵盖全国含油气盆地、全国油气资源地质调查成果、全国油气地球物理勘探、全国页岩气、全国煤层气、全国天然气水合物等数据库)、地球物理数据库(涵盖全国 1:100

万、1:50 万、1:25 万、1:20 万、1:5 万重力、全国区域重力、全国航磁、全国地面磁力、全国电勘查、地震勘查、全国航空放射性、全国测井等数据库)、地球化学数据库(涵盖全国 1:25 万、1:20 万化探、全国多目标地球化学、全国土地质量评价成果等数据库)、遥感调查数据库(全国航空遥感影像、中国资源卫星资料、航天遥感影像、全国矿山环境遥感监测、全国高分卫星等数据库)、钻探数据库(涵盖全国地质钻孔信息、全国重要地质钻孔、中国大陆科学钻探钻岩芯扫描图像库等)、水工环灾害数据库、资料文献数据库、专题专项数据库(涵盖全国矿产资源潜力评价数据库、重要矿产“三率”调查与评价数据库等)、工作管理方面的数据库(涵盖全国探矿权、采矿权、矿业权核查、地质信息元数据库等)。这么多的数据库,一方面还在不断扩充完善,另一方面其本身的实用价值尚未充分体现。然而,绝大多数的成矿规律研究人员实际上不可能拥有上述全部数据,至多也是各自采用自己积累的资料而已。但无论如何,即便是各自积累的资料,无论是数量上还是类型上,也是 10 年前、20 年前所不可比拟的,因此也可以说进入了“相对大数据”时代。比如,1999~2004 年间,“中国成矿体系与区域成矿评价”项目,全国虽然有 202 位矿床学界的专家共同参加,但只掌握 4500 处矿产地的资料(全部矿种);而 2006~2013 年间开展的“全国重要矿产和区域成矿规律研究”,仅仅由矿产资源研究所承担的 16 个矿种(组)所涉及到的矿产地就达 30 600 处。因此,十年间信息增加之快、数据量增加之大是前所未有的。

2.2 地质数据的高速性

地质数据的产生也是非常快的,具有显著的高速性。尤其是遥感地质、航空物探、区域化探等领域甚至连宝玉石矿产资源这样的“小专业”,由于其更多地引进了新技术新方法,其数据都是成倍、甚至呈几何级数地增长。同时,处理速度快,更是“大数据”的时代特点。人们不但需要对数据进行近实时的分析,而且需要以音频、视频、三维图像等迥然不同于传统方式的数据处理技术来表达数据挖掘和加工的结果,没有强大的、快速的处理功能是难以实现的。比如,对地球深部信息的探测,不只是获取地震波反射、折射方面的参数,更需要快速地处理,以便于及时地预测地震是否发生、预报发震的时间、地点、强度等(程陈等, 2013)。这样才能有效避灾,否则就是马后炮了。

2.3 地质大数据的多样性

随着计算机技术的普及,地质大数据从地球物理等传统的需要‘大存储’的‘大专业’扩展到了同位素年代学这样的‘小专业’。从原始信号到原始数据再到图像、声音等,几乎无所不包;而且不同类型之间的数据可以‘互联互通’‘互通有无’;从数据变到图像,再从图像变到数据;‘图文并茂’。专业之间的壁垒也逐渐被打通。比如,成矿系列的研究,以往仅仅是靠专家的脑力劳动来判断矿床与矿床之间是否存在内在联系,矿床与物探异常、与遥感信息之间存在什么样的内在联系,是难以‘主观说清楚’的,因为很少有人既懂得矿床学又精通遥感、物探等专业。“大数据”时代,则可以通过大数据、全数据的采集,在非专业人士主持的情况下也可以分析其间是否存在相互联系——即统计规律;再根据统计规律由专业人士来总结因果联系,即成因规律。不同数据类型之间,也可以分别处理,图像专家处理图像资料,矿床专家分析矿床成因,勘查学家总结勘查经验,殊途同归,共同归纳出成矿规律和勘查规律。

2.4 地质大数据的价值性

对于‘大数据’价值性的理解,研究者们还存在歧义,有的只强调‘价值密度低’(海量数据中真正有用的信息少之又少);有的则强调‘商业价值高’(不用多大投入,仅仅从网络搜集资料就可以带来‘商业价值’)。但一般认为,“价值”指的是“通过分析数据可以得出如何抓住机遇及收获价值”,即‘价值密度低,商业价值高’,地质大数据同样如此。迄今为止,对于大量的物化探异常信息,真正验证了的并不多,找到矿的则更少,但一旦取得找矿突破,其社会经济价值就非常巨大,如西藏的甲玛铜多金属矿床(唐菊兴等,2011)和四川甲基卡外围新发现的锂多金属矿床(王登红等,2013)。

2.5 地质大数据的物质性与非物质性

物质性强调的是地质数据的客观存在性,即具有不以人的意志为转移的物质第一性;而非物质性指的是地质数据又具有被加工处理的可能性。传统的地质学及矿床学都十分强调调查/研究对象的物质第一性。“第一手”“客观真实”“全面”也是地质工作者获取地质数据的基本职业操守。但是,在当今世界‘信息爆炸’的时代,谁也不能保证所有的信息都是真实可信的。那么,除了读者(受众)自身的专业素养(即判断力)之外,利用各种各样的技术来处理数据,去伪存真,去粗取精,成为当今数据技术

领域之主导。维克托·迈尔-舍恩伯格等人在其《大数据时代》一书中也试图将技术与信息、数字化和数据化区分开,“如今的信息技术变革的重点在‘T’上,而不是在‘I’上。现在,是时候把聚光灯打在‘I’上,关注信息本身了”(卞友江,2013)。可见,尽管信息技术指的是处理信息的技术(IT),但信息的物质第一性不应该被忽视,信息本身的内蕴有待于深入挖掘。比如,处理化探异常的技术方法很多,无论是克里格法还是其他方法,都可以得出众多的等值线图。但这些‘异常’及其原始数据本身究竟是代表了成矿还是不成矿,却并非一目了然。因为,成矿元素的分散同样可以形成异常,而元素的分散恰恰是成矿物质聚集的反过程。因此,究竟是处理信息的技术重要还是信息本身更重要,同样是地质大数据所面临的现实问题。显然,在‘大数据’时代,二者不可偏废。没有数据,技术发挥不了作用;没有技术,数据也只能是一堆令人眼花缭乱的数字,难以上升为‘规律’,难以指导人们该采取什么样的实际行动。

2.6 地质大数据的空间性与非空间性

空间性指的是地质数据的空间属性。地质学与数学、化学和物理学不同,与天文学、动物学、植物学则有共同之处,即地质学所研究的对象、所采集的数据往往具有空间属性。不同地区的动物资源、植物资源是不同的,矿产资源也是如此,即‘空间性’。前文中所提到的中国地质调查局建立的一系列数据库,体现的是‘中国地质数据的空间属性’。国外不同国家有不同的地学数据库,中国各个省(甚至更小的行政区)也都在建设数据库(商培林,2015)。众所周知,GIS(地理信息系统)在地球科学领域的发展非常快,其使用范围超出了“地质锤、罗盘、放大镜”的老三件。老三件以采集数据为目的,GIS则是管理数据的高手。可以说,任何一个地质数据都具有空间属性,其价值也就体现在矿产资源分布的空间规律上,因此,在厘定成矿系列、总结成矿规律的过程中,必然要考虑其空间性——即成矿区域特点。毫无疑问,每一个成矿系列都具有空间属性。同时,空间性又具有局限性,在深入研究成矿规律的过程中又要跳出空间的局限性,查明‘成因’,即成矿过程的本质属性,才能更好地指导找矿。比如,长江中下游的矽卡岩型铁铜矿与太行山的矽卡岩型铁铜矿,在空间上并无关联,但在本质上都是‘矽卡岩型’,都跟中酸性侵入岩与碳酸盐岩地层的接触交代成矿作用密不可分,此即非空间性。因此,长江中下游地区的

各种各样的“数据”,在应用到太行山时,就要注意到哪些是具有空间局限性的,哪些是空间无关的,这样才能更好地总结成矿规律并指导找矿。

2.7 地质大数据的时间性与非时间性

地质数据具有“时间”属性,这也是与物理、化学等自然科学不同的。地质学的基础支柱之一就是地质年代表,不同地质时代的岩石、地层、矿床具有不同的分布特征和规律,因而每一个数据也都具有“时代”属性。即便是第四系的土壤样品,在采矿之间与采矿之后、开垦之前与开垦之后也是变化的。同样是矽卡岩型铜矿,长江中下游和冈底斯 2 个成矿带不一样,不只是空间上的差别,而且存在时间上的差别。南岭地区,燕山期的花岗岩与加里东期的花岗岩,都是花岗岩,但在成矿强度、成矿能力、成矿类型等方面明显不同(陈毓川,1983;王登红等,2014b),燕山期是全球知名的重要成矿期而加里东期则不是。这便是“时间性”。当然,不同地质时代的地质样品,又具有共性,比如 30 亿年以前的石灰岩和 3 亿年、0.3 亿年前的石灰岩在地球化学成分和物性参数等方面是有共性的,即“非时间性”。此外,与其他学科一样,大数据的“时间”特色并不局限数据本身的时间属性,还在于获取数据、处理数据的时间效应。比如,现今获取数据比以往更加快捷、高效、等时、无等级差别等。而对于同一个矿床,或者同一个远景区,中国人可以对它进行研究,外国人也可以同时对它进行研究,不一定是近水楼台先得月,即先拿到第一手资料者先出成果。

2.8 地质大数据的因果性与非因果性

传统地质科学的研究,总是想搞清楚“为什么”。比如,板块为什么漂移?恐龙是怎么灭绝的?金刚石为什么来自于地幔?等。但实际上,无论是板块漂移、恐龙灭绝还是金刚石的成因,都尚未取得共识。而且,这些问题从提出之日起,研究成果很多,但似乎并没有突破当初的假说。这是因为,人们始终在各自的研究领域追求着信息的权威性、数据的精确性和证据链的完整性。这种传统的逻辑因果关系的推理方式,正在被“大数据”打破,即“抽样调查”始终无法代替“整体数据”。反之,大数据因为更强调数据的完整性和混乱性,可以帮助我们进一步接近事实真相(卞友江,2013)。也就是说,大数据在不反对“精细”求证的同时,更强调数据的“全部”,哪怕“整体数据”是杂乱无章的,而杂乱无章可能正是事物的本来面目。换句话说,恐龙灭绝的说法多种多

样,“成因复杂”,大家争论不休,而“恐龙灭绝”事实上可能就是由多种多样的原因造成的。板块是怎么碰撞的?假说很多,但谁也不可能用实验来验证(泥巴实验实际上连“抽样调查”都无法比拟),或者可以用肉眼来验证(眼见为实,谁也不可能潜入地幔观察金刚石的形成过程)。因此,大数据时代,人们获取信息,完全可以不需要等待查明原因之后。这对于成矿规律和成矿预测的研究也是深有启发的,即,对于某一个具体矿床的成因争论可能永无止歇,不论是相对简单的“沉积矿产”还是十分复杂的“内生矿产”,都会产生不同看法,但不能阻碍找矿工作,找矿是不必等到查明矿床的成因之后的。当然,掌握规律的找矿才是科学找矿,否则是“碰运气”。

2.9 地质大数据的主体性与非主体性

主体性指的是采集数据、加工数据、发表数据的主体(人以及由人设计的计算机软件等数据数据技术、标准等)对于如何采集数据、如何处理数据、发表哪些数据等具有“先入为主”的特殊性。由于“大数据”给人类的思维带来了变革,即人们由对事物之间因果关系的穷追不舍转向对事物之间相关关系的宏观考察,人类的主观能动性可得到进一步的发挥,数据采集者及读者的主体地位也可以得到进一步的确认。针对同一批数据,仁者见仁,智者见智,甚至“见异思迁”的事例也将层出不穷。比如,当我们看到同一幅地质图时,大地构造学家想到的是“板块碰撞”;勘查地质学家想到的是哪里有矿;地图学家想到的是如何改进“大数据”制图技术(每一幅地质图都是一个“大数据”的产物,既有数据本身,也有相关关系,如断层就是 2 个地质体之间的接触关系);环境保护主义者想到的则可能是哪里还有“世外桃源”;而该图幅的原始制作者(野外地质队员, mapping geologist)想到的可能是当年野外填图的种种逸闻趣事。同样道理,对于成矿规律的研究,即便是面对同一批数据(资料),不同的专家可能得出完全不同的结论。同一个矿床,甲专家可能把它归属于 A 成矿系列,乙专家可能把它归属于 B 成矿系列。数据量越大、数据来源渠道越多,主观性可能越显著。正如 Bill (2013) 所说,“成功分析的关键不是工具和技术本身,使用这些工具和技术的人,才是取得成功的核心要素”。所以,要真正挖掘地质大数据的价值,还需要地学界人士的积极参与,对数据不断地进行交换、整合和解析,才能发现新的地学知识,才能创造新的地学价值。比如,地球化学数据在土地管理方

面的深度融和,有助于“精耕细作”,无疑将推进农业现代化;但地质学家并不十分清楚农作物的生长规律及其对化学元素的选择性,只有地学与农学的共同结合,才能更好、更快地发挥“大数据”的优势。

2.10 地质大数据的客体性与非客体性

客体性(相当于受众性)指的是原先被动地接受数据及其研究成果(或者说信息,包括数据、图片)的任何一个人(或者说“受众”),可以“反客为主”地成为主导者,即被动接受数据的个体(人/计算机等)对于数据可以采取不同的主动的反映,反过来成为成矿规律、成矿系列乃至成矿预测研究的“发动者”。当某人想知道一个自然生长的海蓝宝石的晶体究竟是什么样时,利用“大数据”技术,他马上可以搜集到各种各样的海蓝宝石的图片、数据及文字说明。如果要想投资开发一种“富锂矿泉水资源”新品种,也首先要利用“大数据”查一查有没有这种资源,然后在“去哪里开发、如何开发”等关键问题上作出选择。即,只要你想得到,不怕做不到,而这“一大堆资料”并非哪位科学家事先已经给你准备好了的。那么,谁“想得到”呢?显然,并非每一个人都能“想得到”。因此,“大数据”既是个人的,又是非个人的。显然,专业人士(或者说专家)可能对数据更敏感,其掌握数据、分析数据的能力更强,“收益”也往往更大。

3 “大数据”在成矿规律研究中的应用

在成矿规律研究方面,“大数据”的意义不仅在于掌握庞大的矿产地空间分布和储量、产量等方面的数据信息,更在于对这些数据进行规律性的思辨和客观性的表达,揭示矿产地之所以如此分布的本质原因。也即从地质大数据中充分挖掘有用信息及其规律,通过成矿理论的系统分析,揭示其内在规律,并转化为新的认识或知识,指导未来的地质矿产工作。目前,中国已经开始从大数据的角度对典型矿床、重要矿种的潜力评价开展试点示范,如广东省国土资源档案馆对大宝山矿区地质资料集群化科普服务的示范研究(黄华坤,2015)。

成矿规律是人类对于矿产资源时空分布特征的规律性认识,是主观能动性对于客观存在性的一个认知,认知水平的高低、认知能力的强弱、认知范围的宽窄,都毫无疑问跟“数据量”的大小、数据的类型、数据处理的方式方法有关。因此,要总结好成矿规律,必须充分地掌握资料,这是成矿规律研究的基

本要求。“大数据”技术无疑为大数据时代的成矿规律研究打开了一扇新的大门。由于“大数据”技术常用的各种各样的统计分析技术(如聚类分析、回归分析、空间分析、时间序列分析),实际上在传统的成矿规律、成矿预测研究中也是常用的工具,因此,“大数据”与成矿规律的研究之间并无技术上的阻碍,应该说,“大数据”给成矿规律的研究提供了更加全面的“样本”和更加快捷的技术方法。当然,“大数据”的挖掘,仍然要以人为主导;而成矿规律也是人在研究、总结成矿规律,不是数据本身,计算机也不可能代替“人”,也就是李幼平院士指出的“用‘小数据’指引‘大数据’,牵着互联网的鼻子走”(薛京等,2013)。

与传统的成矿规律研究相比,在“大数据”时代的技术支撑下,成矿规律的研究除了数据量更大、数据类型更多更复杂之外,各种类型数据的集合、集成也是一大特点。这不仅是表达方式的多样化与集成化,原始资料本身也往往是多元化与多样化的,而且不一定是原始资料,或者说不一定从“建立卡片”开始。“大数据”的一个显著特点是包括了大量的非结构化或半结构化数据,如网络日志、图片、音频、视频、地理位置信息等(胡雄伟等,2013),这些信息都是有用的。比如,朱明玉等(2014)在总结中国铬矿的成矿规律并进而探讨深部找矿问题时,就采用了一些重要会议上得到的有关内蒙古贺根山成铬带在2469.89~2474.27 m深度的石油钻孔中见到超镁铁岩、其特征与赫格敖拉矿区含矿超镁铁岩相似的“口头信息”,也采用了严铁雄等铬铁矿专家由个人从业积累得出的一些“经验信息”以及互联网上发布的关于西藏罗布莎深部发现铬铁矿的“电子信息”。

4 “大数据”在成矿系列研究中的应用

矿床的成矿系列,指的是在一定的地质历史时期或构造运动阶段,在一定的地质构造单元及构造部位,与一定的地质成矿作用有关,形成的一组具有成因联系的矿床的自然组合(陈毓川等,2007)。这一概念本身,就明确了只有搞清楚时间、空间、成因及矿床之间的相互联系,才可能具体地厘定出一个成矿系列来。在成矿系列研究方面,“大数据”的应用不仅在于掌握庞大的矿床成因方面的数据信息(包括成矿温度、成矿压力、流体包裹体、同位素、微量元素等矿床地球化学方面的数据),更在于对这些数据从时间、空间、成因和矿床之间的相互内在联系

的角度进行关联性分析,以查明矿床之间的成因联系,查明‘全位成矿’的特点进而为‘缺位找矿’提供依据。

矿床的成矿系列是中国地质科学家程裕淇先生创立的成矿理论,在 1979 发表的《初论矿床的成矿系列问题》一文中,全国只厘定出 19 个成矿系列、7 个亚系列,但充分体现了朴素唯物主义的思想;在 1983 年发表的《再论矿床的成矿系列问题》一文中,则自觉地引入了系统论的思想,这在当时是先进的,但在全国范围内也只建立了 27 个成矿系列、44 个亚系列。到 2006 年陈毓川等发表《三论矿床的成矿系列问题》一文时,历史唯物主义的学术思想已经十分鲜明,大数据的特点也比较显著,不但是矿床成矿系列增加到 214 个、成矿亚系列增加到 434 个,而且新厘定出 38 个成矿系列组和 978 个矿床式。也就是说,当前的矿床成矿系列研究,不再是针对某一区域、某一时代的某一矿床类型,而是一项全局性的工作,也只有掌握全部的资料,“不采取随机分析法(抽样调查)这样的捷径”,而采用所有数据进行分析处理,才能客观、全面地分析矿床与矿床之间的关系,才能厘定成矿系列,才能为成矿预测和地质找矿提供科学依据。如果说传统的‘解剖麻雀’式的典型矿床研究就是‘抽样调查’的话,那么,“采用所有的数据进行分析”,实际上就是‘成矿系列’的研究;而成矿系列的研究,首要的就是查明矿床与矿床之间的联系,并不需要对每一个矿床都进行‘解剖麻雀’。正如维克托·迈尔·舍恩伯格指出的:“大数据时代……最惊人的是,社会需要放弃它对因果关系的渴求,而仅需要关注相关关系。也就是说,只需要知道是什么,不需要知道为什么”。当然,成矿系列的研究不但需要知道是什么,也需要知道为什么,在难以查明一个一个为什么之前,先查明是不是存在‘关联性’则较为现实。这也是当前成矿系列研究的现状。2002 年,王登红等将阿尔泰成矿省的成矿系列厘定为 3 个主要系列、5 个亚系列、8 个矿床式,数量很少,但其基础数据却不包括 883 个地质实体(298 个岩体和 585 个矿床)的名称、地理位置、形态学特征、品位吨位、地质特征和矿床特征等方方面面的资料,还包括了各类地球化学数据约 83 570 个。2007 年,全国范围内一下子厘定出 214 个矿床成矿系列(是阿尔泰成矿省的 71 倍),其中每个矿床成矿系列研究的精细程度显然比不上阿尔泰的 3 个成矿系列。可见,与传统

的矿床成矿系列研究方法不同,大数据时代的成矿系列研究可以先宏观再微观,先把形形色色的具体的矿床之间的相互关系大致框定,而不是要等到把每一个矿床的成因都查明之后。

5 “大数据”在成矿体系研究中的应用

成矿体系是指一定空间范围内各个地质历史时期所形成的矿床及其与成矿作用密切相关的地质要素所共同构成的整体。地质要素包括成矿时代、成矿地质环境、与成矿有关的地质作用及其过程等等。对每一个具体的矿床成矿系列而言,一般涵盖了成矿时代、构造环境、地质成矿作用和成矿元素或矿种这 4 方面的要素,地质构造环境是成矿体系的基础,区域地球化学和区域地球物理是成矿的显示,矿床成矿系列是构成成矿体系的基本单位,成矿区带是成矿体系的载体,而成矿预测则是成矿体系研究的应用目标,通过矿床成矿系列的构筑揭示地壳运动的基本规律、反演地球演化的历史轨迹,则是成矿体系研究的科学目标。因此,采用矿床成矿系列这一基本单位来构筑中国成矿体系比较合适。在成矿体系研究方面,“大数据”的意义不仅在于掌握庞大的地质、地球物理、地球化学、遥感等方面的数据信息,更在于对这些含有意义的数据进行专业化处理,查明通过不同技术手段、不同渠道来源的数据相互之间的内在联系,查明数据与成矿物质来源、运移、聚集、分布、分配的地质过程(王登红等,2011)。在肖克炎等(2006)开发的 MRAS 系统上,陈郑辉等(2011)以《中国成矿体系与区域成矿评价》项目成果的管理和利用为目标,将专家系统与 GIS 地学综合数据库及软件工具相结合;对每一个矿床成矿系列(及亚系列)予以统一编号、详细描述,形成了成矿系列知识库,建立了数字化的成矿系列推理机,实现了大数据与成矿体系的初步结合,可为地质大调查和矿产资源的勘查、管理工作提供便捷的查询服务。同时,也为大数据平台下应用成矿系列和成矿体系理论进行成矿预测创造了条件。

6 “大数据”在矿产预测研究中的应用

在矿产预测研究方面,“大数据”的意义不仅仅在于利用庞大的数据信息和多种多样的数据类型,作出各种各样的异常图件,更在于对这些含有地质

意义的数据进行专业化处理,查明成矿物质分布、分配、富集、贫化的“数字”特点,摸清资源家底和成矿潜力,圈定异常区、远景区,指明找矿方向。同时,在矿产资源的开发利用、保障程度、替代方案、环保安全等方面同样可以根据其“低密度高价值”的特点发挥重要作用,而且往往具有战略意义。

比如,对于矿种的预测是成矿预测的首要工作,社会经济发展的哪个阶段需要什么样的矿产资源,什么样的矿种可以在什么样的场合用得恰到好处,不但是个战略问题,也是个现实问题(即要不要找该矿种的矿床、要不要开发此类矿产资源)。大数据在这方面同样可以大显身手。《大数据》等书中就例举了“手电筒和蛋挞”“啤酒与尿布”“私家车与抗生素”“微博与疫苗接种”等故事。当然,探讨事物间的关联性,在大数据出现之前就已经是普遍采用的工作方法,只是信息量不一样、处理信息的技术也不可同日而语。如,当今社会对于“环保”和“安全”普遍提出了更高的要求,国外一些国家通过借题发挥,大肆渲染“石棉有害论”“石棉恐怖论”“反石棉运动”,要求俄罗斯和中国等国家全面禁止石棉的生产。中国也准备将石棉这一矿种从矿产品名单中扫地出门。那么,石棉为什么如此敏感呢?大数据分析可以给我们以“启发”。中国每年因为车祸而死亡的人数早在十年以前就超过10万了,车祸频发的技术原因之一就是“刹车”,而石棉恰恰是用来制造刹车片和闸瓦的!在驾驶者文明程度和素质远不如国外发达国家的前提下,提高刹车片和闸瓦的摩擦系数,对于中国人的生命安全是多么重要。据国际石棉协会统计,在摩擦材料代用品开发应用的高峰期,卡车交通事故增加了6.4倍,所有机动车辆的交通事故增加了6.2倍。国外之所以鼓吹“反石棉”恰恰是为了推销其石棉替代品,而替代品在替代了石棉行业工人的饭碗的同时,也增加了老百姓行路的风险!对于其他矿种,尤其是石油、煤炭、钢铁、水泥等大宗矿产和锂、铍、铌、钽、稀土元素等新兴产业矿产的预测也要利用“大数据”。

除了对矿种的预测,对于某个地区矿产资源潜力的评价、资源总量的预测以及矿体的定位预测,同样需要“大数据”。比如,中国华南离子吸附型稀土元素,究竟有多大的资源总量和找矿潜力,具体矿体分布在什么地方?除了常规数据资料外,刘新星等(2015)还利用DEM数据提取高程、坡度、坡向、曲率、地形起伏度、地表切割深度、地形特征等各类地

貌因子值,并与已知矿点及矿区进行叠加分析,总结出矿产地所处位置的地貌因子值,即:最佳成矿有利地貌为高程150~500 m、坡度0~20°、地形起伏度100~400 m、地表切割深度40~150 m、地形特征为山顶或山脊,进而划定了最小预测区。对于大数据思维下成矿预测的问题,肖克炎等(2015)有文介绍,此不赘述。显然,大数据相关性分析和数据可视化技术能够让数据“说话”,而且在具体矿产资源预测评价工作中也可以解放劳动力,促进地质找矿突破,但也应该注意是“谁”、需要什么样的条件才可能让数据“说话”,而且说的是真话、有远见的话,不是事后诸葛亮。

7 问题讨论

虽然“信息爆炸”的说法早已有之,但“爆炸”产生的影响效果及社会环境的接受态度是不同的,有的人回避了,有的人被炸“懵”了,有的人不以为然……。如今的大数据浪潮又会产生怎样的社会效益、对科研工作尤其是对成矿规律和成矿系列的研究产生什么样的冲击呢?一方面,成矿规律和成矿系列的研究都需要“全信息”即“全部的数据”,而不是“抽样调查”,要充分地利用大数据;另一方面,数据太多、类型太复杂,千头万绪,又会把人“搞懵”,也要搞一些抽样调查。如何处理这一矛盾?显然,我们不能因为一检索就出现成千上万的文献篇目而如临大敌、甚至产生“反感”情绪以至于刻意反对使用“大数据”、回避“大数据”而是应该充分利用“大数据”时代带来的技术和成果,正确面对“大数据”的挑战,把握“大数据”时代的机遇。“以人为本”,将数据、技术与人结合起来,才能把成矿规律和成矿系列的研究工作深入下去,才能把花花绿绿的“数字地质”转换为切中要害、点石成金的“智慧地质”。应该指出,利用大数据来研究成矿规律和成矿系列,刚刚起步,许多问题尚待探讨,包括:

(1) 信息不对称问题。一般来说,掌握信息越多越主动,但是,这首先要看信息的“含金量”;“垃圾信息”“虚假信息”越多,只能更加背离事物的真相。

(2) 认识不对称问题。管理果园的农场主,看到苹果落地现象的次数要远远多于牛顿,但并没有像牛顿那样总结出“万有引力定律”。可见,由于人的认识能力不同,对“大数据”进行分析判断的能力是不一致的。在1999~2004年间,由全国各个省的

专家归纳出的矿床成矿系列合计达 540 个,比全国汇总组专家厘定的(214 个)要多得多(显然不符合逻辑)。这是由于不同专家认识问题的方法和角度不对称造成的。大数据分析不强调成因,即只要知道是什么而不关注为什么,但‘知道为什么’恰恰是牛顿高明于一般农场主之原因。“一叶知秋”也是同理。因此,在‘信息不对称’的情况下,对于矿床学家尤其是从事‘地质找矿’的矿床学家,更需要掌握成矿理论,知道‘为什么’,才能掌握成矿规律,而不仅仅是搜集资料、建立模型。

(3) 权力不对称问题。无论是个人还是国家,都不希望隐私、情报、信息因为‘大数据’而被泄露。这是由于掌握‘大数据’的个人或集体拥有更大的权力,也更容易得到更多、更大的‘利益’。这也是‘大数据’商业化的理论基础。类似地,基于‘大数据’得出的分析判断或‘规律性’认识也往往被认为更具有权威性。因此,学术权威更需要掌握‘大数据’,一方面可自主地检验既有规律性的认识是否‘放之四海而皆准’,以便于不断地完善,避免主观偏执;另一方面也可以比一般研究人员更快、更好、更权威地服务于社会需要,服务于指导地质找矿。比如,在总结成矿规律时,对于成矿年龄数据的选择就不能因为是 U-Pb 法就采用、K-Ar 法就不采用;在研究成矿系列时,不能只总结典型矿床特别多的中生代的矿床成矿系列,而不管矿产地数量少而地质时代漫长的前寒武纪的矿床成矿系列;在研究成矿体系时,不能只搜集含矿地层、含矿岩体的资料,而不管矿体的底板、顶板等其他要素是否也对成矿起了关键性作用;在成矿预测选择化探资料时,也不能只采用大比例尺的而不用小比例尺的、只看到正异常,而对‘负异常’视而不见。总之,大数据时代,因为有了计算机的辅助,在数据处理方面可以‘省事’,但在结合地质特征进行数据分析时,‘人脑’仍然不能省略,人的权力不能放弃。

(4) “抽样”与‘整体’的关系问题。“大数据”技术要求在数据分析时充分使用全体数据,而不是抽样调查。这看起来很对,问题是‘什么是全体数据’?全体数据从何而来?不‘抽样’又如何得到‘整体’?实际工作中,所谓的‘全体数据’指的是当前能够搜集到的数据的集体,并不代表对所有事物都逐个进行了调查、采样;要得到全部对象的真正的‘全部数据’,只有对每个对象都进行‘抽样’才可能,即首先还是要抽样,而且是‘全部’采样才能得到全部数据。

因此,也有人认为‘大数据’是个伪命题,只是个相对概念。比如说,要研究斑岩铜矿的成矿规律,传统的办法是对不同时代、不同地区、不同构造环境下形成的一些典型矿床的共同特征进行归纳总结,而大数据情形下,需要更多地使用‘全体数据’,但这些‘全体数据’究竟是什么,代表了什么样的斑岩铜矿,代表性如何,同样是需要矿床学家加以辨析的。否则,即便是采集到了成千上万个斑岩铜矿的资料(假定为‘全体数据’),假如其中中国的数据一个也没有,则不能得出全球斑岩铜矿的成矿规律,因为中国也是拥有众多斑岩铜矿的国家之一。可见,数据的搜集是否达到‘全体’,也需要‘抽样调查’来审查。

(5) 第一手资料和第 n 手资料的问题。打开一些学术刊物的网页,常常出现文章被下载或引用或点击的‘排行榜’,一些重要文章被彰显在榜首,但一细看就发现问题,第一位与第二位的‘引用率’、“下载率”常常差几十倍,而被点击、下载的次数可达到几万、几十万次,意味着所有地质工作者人均下载一次还不止,而真正被阅读的纸质文章并不在‘点击率’、“下载率”统计范围内。因此,在‘大数据’时代,第一手资料往往被忽视,“转引自”“据……”的资料被大量的、累次地使用,虚假资料也仍然可以被引用了又引用。这也是‘大数据’时代的特点,即数据的‘含金量’越来越低(价值密度低),是需要我们警惕的。

8 结 语

总之,‘大数据’时代为全国乃至全球成矿系列的厘定创造了条件,但既要充分地搜集资料,又要发挥人的主观能动性,不能偏废;‘大数据’成矿规律研究的战略意义不仅仅在于掌握庞大的数据信息,更在于对这些含有地质意义的数据进行各个地质专业乃至跨专业的专业化处理,使之更有针对性,厘定成矿系列更加准确,总结成矿规律更加全面客观;“大数据”成矿预测工作也不仅仅是一个物探、化探、遥感等数据的采集过程和圈等值线的作图过程,同样是一个数据分析、数据转化与数据验证的工作。成矿规律、成矿系列和成矿体系的研究,可以看成是地质大数据的采集与转化过程,而成矿预测就是一个验证过程。大数据时代,专家的作用不是可有可无的,而是更加关键,否则事物的真相就会被海量的数据‘稀释’掉,甚至得到错误的结论。

References

- Bian Y J. 2013. The research of the conception of Big Data [J]. News Study Tribune, 14(5) : 25-28 (in Chinese).
- Bill F. 2013. Taming the Big Data tidal wave : Finding opportunities in huge data streams with advanced analytics [M]. In : Huang H, et al. Beijing : People's Posts and Telecommunications Publishing House. (in Chinese).
- Chen Y C. 1983. The metallogenetic series of the rare-earth, rare and nonferrous metal deposits related to the Yenshanian granites in South China [J]. Mineral Deposits, 2 : 15-24 (in Chinese with English abstract).
- Chen Y C, Li T D and Peng Q M. 1999. Mineral resources and the sustainable development [M]. Beijing : China Science & Technology Press. 1-148 (in Chinese).
- Chen Y C, Pei R F and Wang D H. 2006. On minerogenetic (metallo-genetic) series third discussion [J]. Acta Geologica Sinica, 10 : 1501-1508 (in Chinese with English abstract).
- Chen Y C, Wang D H, Zhu Y S, Xu Z G, Ren J S, Zhai Y S, Chang Y F, Tang Z L, Pei R F, Teng J W, et al. 2007. Chinese mineralization system and regional metallogenetic evaluation [M]. Beijing : Geological Publishing House (in Chinese).
- Chen Z H, Wang D H, Chen Y C, Ying L J and Fu X J. 2011. Preliminary implementation of the expert system of metallogenetic series of ore deposits in China [J]. Acta Geoscientia Sinica, 32(4) : 397-404 (in Chinese with English abstract).
- Cheng C and Shi W B. 2013. Big data mining analysis application in earthquake research [J]. Information System Engineering, 25(12) : 27-28 (in Chinese).
- Cheng Y Q, Chen Y C, and Zhao Y M. 1979. Preliminary discussion on the problems of minerogenetic series of Mineral Deposits [J]. Bulletin Chinese Acad. Geo. Sci., 2 : 32-58 (in Chinese with English abstract).
- Cheng Y Q, Chen Y C, Zhao Y M and Song T R. 1983. Further discussion on the problems of minerogenetic series of mineral deposits [J]. Bulletin Chinese Acad. Geo. Sci., 2, 1-64 : 134-135 (in Chinese with English abstract).
- Hu X W, Zhang B L and Li D F. 2013. Overview of Big Data research and application (part A) [J]. Standard Science, 50(9) : 29-34 (in Chinese).
- Huang H K. 2015. Discussion on science popularization application mode of geological data under the support of big data [J]. Lantai World, 30(5) : 124-125 (in Chinese).
- Li G J and Cheng X Q. 2012. Research status and scientific thinking of Big Data [J]. Bulletin of Chinese Academy of Sciences, (6) : 647-657 (in Chinese).
- Liu X X, Chen Y C, Wang D H, Huang F and Zhao Z. 2015. The metallogenetic geomorphic condition analysis of the ion-absorbing type rare earths ore in the eastern Nanling region based on DEM data [J]. Acta Geoscientia Sinica, in press (in Chinese with English abstract).
- Ma G Z, Ni Y F and Gong X P. 1982. Mathematical geology application in metallogenetic prospective prediction in Luqu-Pingwu region [J]. Acta Geologica Sichuan, 3(2) : 114-115 (in Chinese).
- Mayer-Schönberger V and Cukier K. 2013. Big Data : A revolution that will transform how we live, work, and think [M]. New York : Houghton Mifflin Harcourt Publishing Company. 1-158.
- Niu W J and Yang H. 2001. Study of universal Kriging interpolation for super-giant data volume [J]. Oil Geophysical Prospecting, 36(6) : 672-679 (in Chinese with English abstract).
- Shang P L. 2015. The thinking of Shanxi geological information development in Big Data era [J]. Huabei Land and Resources, 14(3) : 105-107 (in Chinese).
- Tang J X and Wang D H. 2011. Jiamia in Tibet, et al. super large copper polymetallic deposits exploration and research progress [J]. Mineral Deposits, 30(2) : 177-179 (in Chinese).
- Toffler A. 1981. The third wave [M]. New York : Bantam books.
- Wang D H, Chen Y C, Xu Z G, Li T D and Fu X J. 2002. The metallogenetic series and metallogenetic regularity study of Altay Metallogenetic Province [M]. Beijing : Atomic Energy Press. 493 (in Chinese).
- Wang D H, Chen Y C, Xu Z G, Chen Z H, Shen B F, Tang X L and Pei R F. 2011. Advance in the study of mineralization system and its application to assessment of mineral resources [J]. Acta Geoscientia Sinica, 32(4) : 385-395 (in Chinese with English abstract).
- Wang D H and Fu X F. 2013. Lithium ore prospecting breakthrough in Jiayika peripheral region of Sichuan province [J]. Rock and Mineral Analysis, 32(6) : 987 (in Chinese).
- Wang D H, Wang C H, Xing S W, et al. 2014a. The Record of mineral geology of China-the list volume of the place name of mineral deposits [M]. Beijing : Geological Publishing House (in Chinese).
- Wang D H, Chen Z Y, Huang F, Wang C H, Zhao Z, Chen Z H, Zhao Z and Liu X X. 2014b. Discussion on metallogenetic specialization of the magmatic rocks and related issues in the Nanling region [J]. Geotectonica et Metallogenia, 38(2) : 230-238 (in Chinese with English abstract).

English abstract).

Xiao K Y, Wang Y Y and Chen Z H. 2006. The new technology and new model of Chinese mineral resources evaluation [M]. Beijing: Geological Publishing House (in Chinese).

Xiao K Y, Sun L, Li N, Wang K, Fan J F and Ding J H. 2015. Mineral resources assessment under the thought of big data [J]. Geological Bulletin of China, 34(7): 1266-1272.

Xue J and Xu Y. 2013. The simple UCL concept: To visit the Chinese academy of engineering Li Youping [J]. TV technology, 37(12): 11-12 (in Chinese).

Zhu M Y, Wang D H, Li L X, Liu X X, Zhang J, Zhou X, et al. 2014. The metallogenic regularities of china chrome ore [M]. Beijing: Geological Publishing House (in Chinese).

附中文参考文献

Bill F. 2013. 驾驭大数据 [M]. 黄海, 等译. 北京: 人民邮电出版社.

卞友江. 2013. “大数据”概念考辨 [J]. 新闻研究导刊, 14(5): 25-28.

陈毓川. 1983. 华南与燕山期花岗岩有关的稀土、稀有、有色金属矿床成矿系列 [J]. 矿床地质, 2: 15-24.

陈毓川, 李廷栋, 彭齐鸣. 1999. 矿产资源与可持续发展 [M]. 北京: 中国科学技术出版社. 1-148.

陈毓川, 裴荣富, 王登红. 2006. 三论矿床的成矿系列问题 [J]. 地质学报, 10: 1501-1508.

陈毓川, 王登红, 朱裕生, 徐志刚, 任记舜, 翟裕生, 常印佛, 汤中立, 裴荣富, 滕吉文, 等. 2007. 成矿系列体系与区域成矿评价 [M]. 北京: 地质出版社.

陈郑辉, 王登红, 陈毓川, 应立娟, 傅旭杰. 2011. 中国成矿体系专家系统的初步实现 [J]. 地球学报, 32(4): 397-404.

程陈, 史文博. 2013. 大数据挖掘分析在地震科研中的应用 [J]. 信息系统工程, 25(12): 27-28.

程裕淇, 陈毓川, 赵一鸣. 1979. 初论矿床的成矿系列问题 [J]. 中国地质科学院院报, 1: 32-58.

程裕淇, 陈毓川, 赵一鸣, 宋天锐. 1983. 再论矿床的成矿系列问题 [J]. 中国地质科学院院报, 2: 1-64, 134-135.

胡雄伟, 张宝林, 李抵飞. 2013. 大数据研究与应用综述(上) [J]. 标

准科学, 50(9): 29-34.

黄华坤. 2015. 大数据支持下的地质资料档案科普应用模式探讨 [J]. 兰台世界, 30(5): 124-125.

李国杰, 程学旗. 2012. 大数据研究: 未来科技及经济社会发展的重大战略领域——大数据的研究现状与科学思考 [J]. 中国科学院院刊, (6): 647-657.

刘新星, 陈毓川, 王登红, 黄凡, 赵芝. 2015. 基于 DEM 的南岭东段风化壳离子吸附型稀土矿成矿地貌条件分析 [J]. 地球学报, 待刊.

马光中, 倪彦福, 龚学屏. 1982. 数学地质在碌曲—平武一带作成矿远景预测的应用 [J]. 四川地质学报, 3(2): 114-115.

牛文杰, 杨钦. 2001. 超大数据体泛克里金插值的研究 [J]. 石油地球物理勘探, 36(6): 672-679.

商培林. 2015. 大数据时代山西地质信息化发展思考 [J]. 华北国土资源, 14(3): 105-107.

唐菊兴, 王登红. 2011. 西藏甲玛等超大型铜多金属矿床勘查与研究新进展 [J]. 矿床地质, 30(2): 177-179.

王登红, 陈毓川, 徐志刚, 李天德, 傅旭杰. 2002. 阿尔泰成矿省的成矿系列及成矿规律研究 [M]. 北京: 原子能出版社. 493.

王登红, 陈毓川, 徐志刚, 陈郑辉, 沈保丰, 汤中立, 裴荣富. 2011. 成矿体系的研究进展及其在成矿预测中的应用 [J]. 地球学报, 32(4): 385-395.

王登红, 付小方. 2013. 四川甲基卡外围锂矿找矿取得突破 [J]. 岩矿测试, 32(6): 987.

王登红, 王成辉, 邢树文, 等. 2014a. 中国矿产地质志·矿产地名录卷 [M]. 北京: 地质出版社.

王登红, 陈振宇, 黄凡, 王成辉, 赵芝, 陈郑辉, 赵正, 刘新星. 2014b. 南岭岩浆岩成矿专属性及相关问题探讨 [J]. 大地构造与成矿学, 38(2): 230-238.

肖克炎, 王勇毅, 陈郑辉. 2006. 中国矿产资源评价新技术与评价新模型 [M]. 北京: 地质出版社.

肖克炎, 孙莉, 李楠, 王琨, 范建福, 丁建华. 2015. 大数据思维下的矿产资源评价 [J]. 地质通报, 34(07): 1266-1272.

薛京, 许盈. 2013. 大道至简的 UCL 理念——访中国工程院院士李幼平先生 [J]. 电视技术, 37(12): 11-12.

朱明玉, 王登红, 李立兴, 刘新星, 张建, 周详, 等. 2014. 中国铬矿成矿规律 [M]. 北京: 地质出版社. 1-169.