



## 电池材料数据库的发展与应用

吴思远 王宇琦 肖睿娟 陈立泉

### Development and application of battery materials database

Wu Si-Yuan Wang Yu-Qi Xiao Rui-Juan Chen Li-Quan

引用信息 Citation: *Acta Physica Sinica*, 69, 226104 (2020) DOI: 10.7498/aps.69.20201542

在线阅读 View online: <https://doi.org/10.7498/aps.69.20201542>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

基于材料基因组方法的锂电池新材料开发

Development of new lithium battery materials by material genome initiative

物理学报. 2018, 67(12): 128801 <https://doi.org/10.7498/aps.67.20180657>

非晶合金的高通量制备与表征

Combinatorial fabrication and high-throughput characterization of metallic glasses

物理学报. 2017, 66(17): 176106 <https://doi.org/10.7498/aps.66.176106>

基于数据库进行乏燃料鉴别的多元统计分析研究

Identification of spent nuclear fuel with multivariate analysis based on database

物理学报. 2019, 68(9): 090204 <https://doi.org/10.7498/aps.68.20190107>

Li离子电池负极材料石墨炔在B, N掺杂调控下的储Li性能优化

Optimized Li storage performance of B, N doped graphyne as Li-ion battery anode materials

物理学报. 2019, 68(21): 213601 <https://doi.org/10.7498/aps.68.20191161>

太阳能电池材料缺陷的理论与计算研究

Theoretical and computational study on defects of solar cell materials

物理学报. 2020, 69(17): 177101 <https://doi.org/10.7498/aps.69.20200656>

高温压电材料、器件与应用

Review of high temperature piezoelectric materials, devices, and applications

物理学报. 2018, 67(20): 207701 <https://doi.org/10.7498/aps.67.20181091>

专题：固态电池中的物理问题

## 电池材料数据库的发展与应用\*

吴思远<sup>1)2)</sup> 王宇琦<sup>1)2)</sup> 肖睿娟<sup>1)2)†</sup> 陈立泉<sup>1)2)</sup>

1) (中国科学院物理研究所, 清洁能源实验室, 北京 100190)

2) (中国科学院大学, 北京 100049)

(2020年9月17日收到; 2020年10月21日收到修改稿)

基于自动化技术和计算机技术的高通量方法可快速提供数以万计的科研数据, 对如何科学、高效的管理科研数据提出了新的挑战. 可充放的二次电池作为一种清洁高效的能源存储器件, 是电动汽车发展的关键, 也是风/光电储能的首选. 电池器件性能的提升与电池新材料的研发密切相关, 电池材料数据库的发展可在电池材料研发中引入基于大数据的新兴方法, 加速电池材料的开发. 本文从电池材料数据的获取、通用性及特定性质的电池材料数据库构建、大数据方法对电池材料研发的促进和发展电池材料数据库所面临的挑战等方面对电池材料数据库的发展和应用进行了介绍.

**关键词:** 电池材料, 高通量计算, 数据库, 材料基因**PACS:** 61.68.+n, 66.10.Ed, 82.47.Aa**DOI:** 10.7498/aps.69.20201542

## 1 引言

在材料研发的诸多领域中, 锂二次电池材料的开发对于能源的清洁高效利用及环境的可持续发展十分重要<sup>[1]</sup>. 当前锂电池产业面临的关键问题是开发出全新的锂电池材料以提升下一代锂二次电池的能量密度、功率密度和安全性能<sup>[2]</sup>. 近年来“材料基因组”作为一种新的研究方法, 有效地加速了材料从研究到应用的进程, 降低了材料研发的成本<sup>[3,4]</sup>. 基于材料基因组思想的高通量技术为科学研究提供了大量的数据, 也对如何高效、完备的管理和使用科研数据提出了挑战, 建立高通量材料研究的相关数据库将有效加速新材料的探索和研发, 并将为材料研究领域中引入机器学习、数据挖掘等人工智能技术奠定基础. 目前材料数据库的研究主要涉及材料数据的产生、归类和应用三个方面, 如图1所示. 数据的来源包括实验数据和计算数据,

除了文献和已有资料中可收集的大量数据外, 高通量实验和高通量计算也提供了越来越可观的数据; 对于收集的数据通常需要根据其获得方式、精确程度和所关联的物性等进行归类, 为数据匹配相应的标签, 以实现数据库的建立; 从构建数据库的数据中, 可以根据应用需求直接筛选符合条件的材料, 也可以通过大量的材料学数据借助机器学习算法来挖掘材料宏观性质与微观结构之间的关联. 随着大量材料数据的出现和人工智能算法的优化, 材料数据库在未来协助研究人员优化和设计电池新材料方面将发挥越来越显著的作用<sup>[5-7]</sup>.

图2显示了几种具有代表性的材料数据库的出现时间. 这其中以按照材料类型建立的数据库为主, 例如为具有某种共同用途或具有某种共同结构特征的材料建立的数据库, 包含每种材料的多种物理、化学等性质<sup>[8-10]</sup>; 也有以某种特定材料性质建立的数据库, 例如针对离子在固体中传输性质的数据库<sup>[11,12]</sup>. 早期的材料数据库常常是针对一类材料

\* 国家重点研发计划 (批准号: 2017YFB0701600) 和国家自然科学基金 (批准号: 51772321) 资助的课题.

† 通信作者. E-mail: [rjxiao@iphy.ac.cn](mailto:rjxiao@iphy.ac.cn)

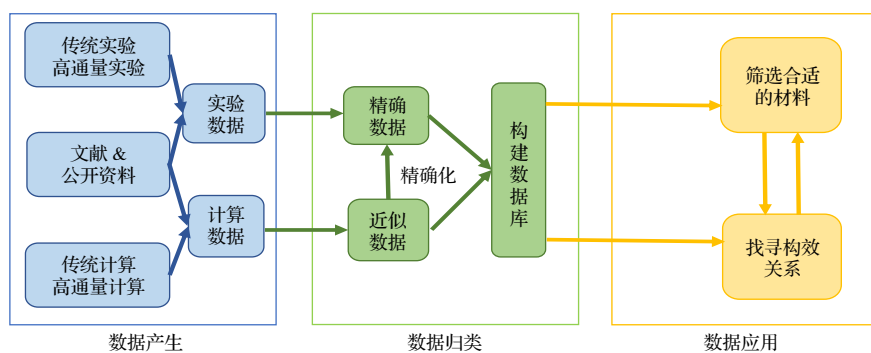


图 1 材料数据的产生、归类和应用程序

Fig. 1. Flowchart of creation, classification and application of materials data.

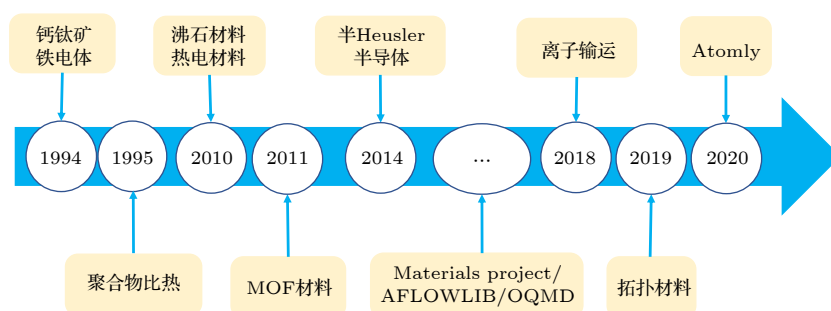


图 2 各类材料数据库的出现时间

Fig. 2. Appearance time of various materials databases.

的一个具体性质, 例如钙钛矿结构铁电材料的声子性质数据<sup>[13]</sup>, 非线性光学材料的阴离子基团性质等<sup>[14]</sup>; 在 2000 后, 逐步出现了热电材料、半 Heusler 半导体、沸石材料和拓扑材料等各类材料数据库<sup>[15–18]</sup>; 近十年来, 随着自动化高通量计算的发展, 目前已实现了对无机晶体数据库中大量已知结构化合物的计算, 因此出现了 Materials Project, AFLOW, OQMD, Atomly 等多个包含各种可计算物性的通用型材料计算数据库<sup>[19–22]</sup>.

可充放的二次电池作为下一代储能器件的首选, 受到了人们的广泛关注. 建立数据信息丰富的电池材料数据库将有助于研究人员从数据获取、数据挖掘和数据预测各个阶段实现电池新材料的探索. 正确认识材料结构与性能之间的关系, 可以合理的筛选、优化和设计新材料, 进而加速材料从研发到应用的过程, 降低材料的开发成本. 本文将在第二部分介绍电池材料数据的来源及通用型和特定性质型的电池材料数据库; 第三部分介绍目前使用电池数据进行材料筛选和机器学习的进展; 第四部分介绍发展电池材料数据库所面临的挑战.

## 2 电池材料数据库的构建

### 2.1 电池材料数据的获取

与图 1 所展示的各种材料数据的获取方式相似, 电池材料的数据也主要来源于实验和计算两个方面. 实验数据的收集和整理主要来源于已发表各类文献, Ghadbeigi 等<sup>[23]</sup>从科技文献中手工收集了大量电池材料数据并构建了数据库. 计算机技术的发展, 特别是基于自然语言的文本挖掘功能的实现, 显著加快了从已发表文献中获取实验数据的自动化进程. Huang 和 Cole<sup>[24]</sup>采用自行编写的 ChemDataExtractor 建立了从文献中自动收集电池数据的方案, 构建了总条目 29 万余条, 包含容量、电导率、库伦效率、能量密度和电压共五种性质的数据库.

理论模拟也为电池材料提供了丰富的数据集. 电池的模拟包括在原子分子尺度、微观尺度和器件在宏观尺度上电池各类性质进行模拟. 如分别采用原子尺度的 DFT 和 DFTB 计算从电荷转移角度以及采用有限元的微观尺度从 Li 浓度梯度引发的应力应变等来阐述界面问题<sup>[25,26]</sup>. 近年来快速发展的高通量计算主要是基于密度泛函理论的高通量计算, 通过设计一系列运算流程, 实现对材料原子

尺度本征性质的大批量自动化计算. 材料中各种不同的物性会涉及到不同的计算方法, 表 1 列出了目前已实现高通量化计算的各种材料性质. 针对电池材料的反应机制, 还可通过热力学数据获得各种材料的理论能量密度, 为实际分析和筛选电极材料提供参考. Peng 等<sup>[27]</sup>与 Zu 和 Li<sup>[28]</sup>分析了过去 60 年电池能量密度的增长趋势并计算了不同体系 Li, Na, Mg, Al 和 Zn 电池的理论能量密度; Wu 等<sup>[29]</sup>使用电池材料的基本数据计算了不同 18650 电芯的实际能量密度; Wang 等<sup>[30]</sup>计算了不含锂的正极材料的嵌锂性能, 得到其理论能量密度; Cao 等<sup>[31]</sup>收集了理论能量密度高的材料的热力学数据, 用于寻找高能量密度电极.

表 1 高通量计算所能获得的材料性质  
Table 1. Properties achieved by high-throughput calculations.

计算数据	物化性质	材料种类
总能量	相图、反应路径、形成能	热力学稳定材料
电子结构	带隙、电子传输、电荷分布	特定电学性质材料
原子磁矩	磁构型、磁矩、磁阻等	磁性材料
声子谱	晶格振动、红外吸收谱	动力学稳定材料
力学模量	弹性模量、泊松比等	力学材料
复数介电常数	介电性质	介电材料
反射系数	反射/吸收率	光学材料
吸附能/位置	表面吸附过程	材料表面设计
晶格匹配	界面力学/界面化学稳定性	材料界面设计
离子扩散	离子迁移路径、势垒等	离子导体

在数据的获取过程中, 需要关注数据产生的条件和数据的误差<sup>[32,33]</sup>. 对于实验数据, 测量环境(如温度、压力等)和测量方法常常会影响数值的大小, 那么后续的数据挖掘则需要对数据进行归类, 在相同条件下测量的数据间可以进行更为科学的比较. 对于理论模拟的数据, 设定相同模拟参数则较为容易, 例如在基于密度泛函的高通量计算中, 通过设定相同的关联函数、积分密度和收敛条件等参数, 可以将数据的准确度控制在相同的范围. 实验数据与计算数据相结合的数据库构建思想目前获得了广泛的认同<sup>[34,35]</sup>. 数据类型的全面和准确是进一步对电池材料数据进行大规模分析和挖掘的基础.

## 2.2 典型的材料计算数据库

电池材料的诸多性质中, 脱/嵌锂电位、热力学稳定性和化学稳定性等均可从密度泛函计算得

到的能量、电子结构等信息中获得<sup>[36]</sup>, 因此包含高通量密度泛函计算结果的通用型材料数据库都可用于电池材料本征性质的研究. 国际上已有多个研究团队推出了包含体系能量、能带结构、力学模量和热力学相图等信息的材料数据库, 其材料结构的来源既包括无机晶体数据库中已有物质, 也包括大量由已有物质衍生出的虚拟结构, 为发现新材料提供了条件. 表 2 列出了几种公开的通用型材料数据库及用于构建该数据库的高通量计算软件. 其中 Materials Project 数据库除了收录密度泛函的计算数据外, 还开发了将基本计算数据转化为电池性质数据的模块, 可获得电压曲线、理论容量、不同锂化学势下的稳定性等用于电池材料研发的数据, 可通过 Battery Explorer 模块进行查找. Atomly 数据库是中国原创的材料数据库, 包含 14 万余种材料的电子结构信息和 4 万余组热力学相图信息, 且含有通过机器学习获得的势函数, 可有效加速分子动力学的模拟进程, 有望为电池材料提供大量动力学方面的研究数据<sup>[22]</sup>.

表 2 国内外典型的通用型计算材料数据库及公开发布的高通量计算软件<sup>[19–22]</sup>

Table 2. Typical database for computational materials<sup>[19–22]</sup>.

数据库名称	高通量计算软件
Materials Project	Pymatgen
AFLOWLIB	AFLOW
OQMD	OQMD
Atomly	—

除了从通用型的材料数据库中获取电池材料信息外, 还有为电池材料某一特定性质构建的数据库, 其中以几何和半经验方法计算得到的锂离子输运动力学数据库为主, 包括我们在 2018 年推出的电池材料离子输运数据库<sup>[11]</sup>和上海大学 2020 年上线的离子传输特征数据库<sup>[12,37]</sup>. 电池材料的离子输运性质与电池器件的充放电速率密切相关, 也是开发新型固体电解质的主要指标之一. 实验中常通过电化学阻抗谱或核磁共振光谱来获取材料的离子传输信息, 理论方法对离子输运现象的模拟则经历了由晶体中几何空间进行预估<sup>[37–39]</sup>、通过半经验势函数进行估算<sup>[40]</sup>和采用基于密度泛函的过渡态方法精确计算<sup>[41]</sup>的几个阶段. 精确计算所需的计算量较大, 为了在初始阶段实现大规模的材料筛选, 基于半经验势函数的键价方法由于能给出离子

运输势垒的变化趋势, 因此被用来作为快离子导体初筛的方法之一. 我们用高通量键价计算的结果构建了电池材料离子输运性质数据库<sup>[11]</sup>. 如图 3(a) 所示, 该数据库包含了采用键价方法计算得到的 21204 种无机晶体化合物中的离子输运势垒, 其中包括含 Li 的化合物 4535 种, 含 Na 的化合物 4344 种, 含 K 的化合物 2808 种, 含 Mg 的化合物 2145 种, 含 Zn 的化合物 2180 种、含 Al 的化合物 5192 种. 目前该数据库具备三种便捷的查询方式, 包括根据化合物的元素组成进行查询、根据化学式进行查询、根据离子输运类型及离子迁移势垒的数值范围进行查询. 利用该数据库可快速排除已知结构化合物中离子迁移势垒较高的物质, 为进一步探寻快离子导体有效地缩小了范围. 同时, 如图 3(b) 所示, 数据库所包含的大量化合物中, 不仅有迁移势垒小的结构, 也有迁移势垒大的结构, 这为后续的数据挖掘和机器学习提供了完备的样本集. 上海大学施思齐研究组<sup>[37]</sup> 则采用几何分析的方法, 利用 Voroni 多边形镶嵌模型寻找扩散路径并编写了 CAVD 程序, 为进一步使用第一性原理 NEB 计算势垒构建了初始输入文件<sup>[39]</sup>.

### 3 电池材料数据库对材料研究的促进

#### 3.1 利用材料数据库进行新材料的筛选设计

材料数据库的建立可以帮助我们加深对已有材料的理解, 发现具有目标物性的新材料. 当人们对某一性质所对应的原子结构或电子结构特征已有清晰认识时, 可以从数据库中直接寻找具有这一特征的化合物. 例如: 在确认非线性光学材料的性质与阴离子基团结构的关联后, Avdeev 等<sup>[14]</sup> 通过寻找具有特定阴离子点群特征的结构来寻找新的非线性光学材料; 在发现了电子结构特征与材料拓扑性质的关联后, Zhang 等<sup>[18]</sup> 发现了数千种新的拓扑材料. 另一种筛选方式是直接计算出目标物性, 选出达到应用要求的材料, 电池材料的筛选大多使用这种直接筛选的方式. 例如 Kirklin 等<sup>[42]</sup> 从 515 种硅化物、锡化物和磷化物中以电化学势、体积变化和容量为标准筛选出  $\text{CoSi}_2$ ,  $\text{TiP}$ ,  $\text{NiSi}_2$  等几种性能优于石墨的负极材料; Zhu 等<sup>[43]</sup> 以锂电势和热力学稳定性为标准筛选出对金属锂负极稳定的化合物; Wang 等<sup>[44]</sup> 以离子输运势垒为标准筛选出可以提高  $\text{Li}_3\text{PS}_4$  离子电导率的氧掺杂和锌氧共掺杂方案.

#### 3.2 利用材料数据库进行构效关系的挖掘

随着高通量计算和高通量实验带来的材料大数据的出现, 机器学习成为探索材料的微观结构与宏观性质之间关联的新方法. 使用机器学习方法探究材料中的构效关系, 是借助数据挖掘算法在所关注的目标物性与材料的组分、结构等变量间建立映射关系. 如图 4 所示, 对于电池材料而言, 目标物性可以是嵌锂电位、电子电导、脱嵌锂体积变化、离子迁移势垒等各种为满足应用需求所要达到的物性; 用于描述材料的组分、结构的变量称为描述因子, 研究人员可以根据对于材料的认识来进行构建, 例如用晶格参数、对称性等描述其晶体构造, 用配位数、键长、键角等描述局域化学环境等. Sendek 等<sup>[45]</sup> 通过选择锂离子周围配位数、配位距离等描述符并利用机器学习中的多元线性回归算法判断各种晶体结构作为锂电池固态电解质材料的可能性; Liu 等<sup>[46]</sup> 利用支持向量机算法, 探究了掺杂元素化合价、掺杂离子半径、掺杂元素泡利电负性等描述因子与锂电池/固态电解质界面稳定性之间的关系. 表 3 中列举了采用机器学习方法研究

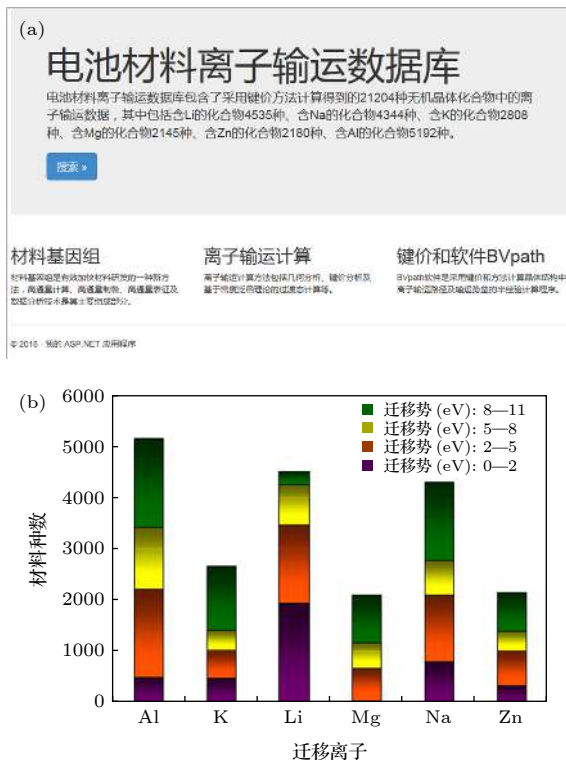


图 3 (a) 电池材料离子输运数据库网站页面; (b) 数据种类  
Fig. 3. (a) The database of ion transport properties for battery materials; (b) data distributions for various types of materials.



图 4 数据挖掘方法在探究材料构效关系中的应用

Fig. 4. Data mining method applied in exploring the relationship between structure and properties.

表 3 机器学习模型应用于二次电池的构效关系  
Table 3. Application of machine learning method in the research of secondary batteries.

关心的问题	输入量(描述符)	输出量(目标性质)
固态电解质	原子结构以及它们的XRD图像信息、化学键数目、子晶格化学键的离子性、原子配位数、键长、位能、熔点、沸点...	离子电导率或离子迁移能
聚合物电解质	化学结构、组成比率、处理温度、Mordred描述因子...	离子电导率
锂电电极	化学键、原子半径、单位原子体积、质量密度、子晶格电负性、Li原子周围原子数变化...	热力学相稳定性
界面	热力学稳定性、结构和动力学参数...	界面态
电池制造	活性材料质量比率、粘性、固液比率...	孔隙率和电极的质量负载

二次电池中各类构效关系的实例,可以看出,这种基于大数据的分析方法可广泛应用于固态电解质、聚合物电解质、电极/电解质界面和电池制造等各方面的研究.

## 4 发展电池材料数据库的挑战

### 4.1 建立面向应用的电池材料数据库

电池材料数据库需面向各类科研及工业开发,因此在数据获取、数据管理和数据使用方面要兼顾多种应用场景.图5列举了电池材料数据库在构建过程中需要优化的各个方面.数据的获取无论是计算数据还是实验数据,都需要关注数据的误差范围

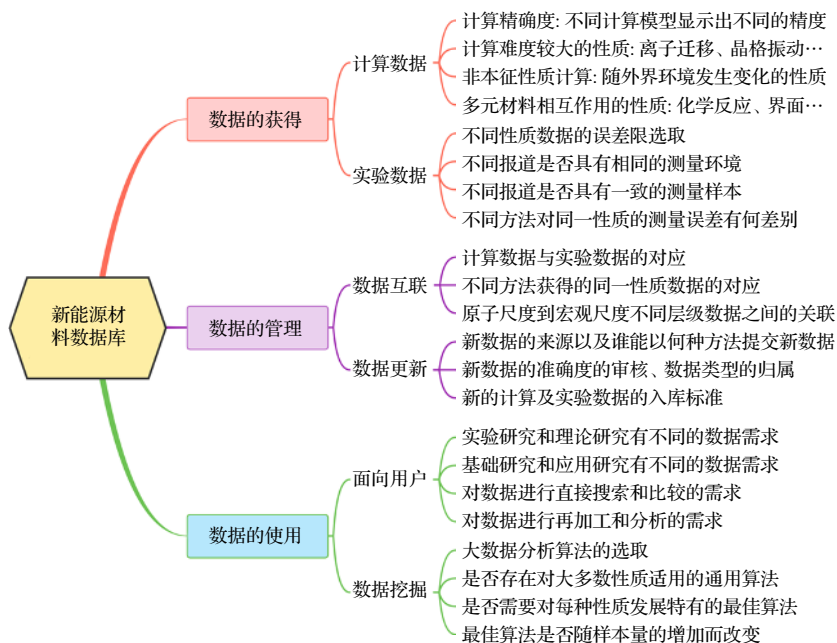


图 5 新能源材料数据库的主要技术挑战

Fig. 5. The main technologic challenges in the development of energy materials database.

表 4 不同层级的数据库类型、用途和使用方法

Table 4. TType, application and usage of battery materials database in various scales.

层级	数据库类型	用途	使用方法
原子尺度	基于理想材料模型获得的材料在原子尺度的本征性质数据	了解所选用材料本身所具备的性质特征	查询及挖掘原子结构及对应的电子结构、离子输运势垒等数据,帮助寻找到具有目标物性的材料
微观尺度	引入实际材料中的缺陷和微观构型后获得的实际材料性质数据	了解微观结构对材料性质的调制	查询及挖掘缺陷、粒径大小、颗粒形状、比表面积等一系列变量描述下的材料性质数据,帮助实现对所选材料的性质改善
外场效应	随电场、温度等外场条件改变时获得的材料性质数据	了解材料性质对外界环境的响应	查询及挖掘材料性质数据随外场条件的变化函数,帮助设计电化学稳定的电池材料
多相作用	将单一材料性质数据扩展到多种材料之间相互作用的性质数据	了解界面等由多相作用所决定的性质数据	查询及挖掘电池中界面的组分、性质数据,帮助选取相匹配的组成电池的各种材料
宏观尺度	电池器件的性能数据及充放电过程中电池材料的性质数据	实现材料性质数据与电池器件性能的关联	查询及挖掘上述四层性质数据与电池器件性能之间的联系,帮助实现从材料到电池的整体设计

和获取条件. 对于数据管理, 需考虑不同方法所获得的数据之间如何对应、不同空间或时间尺度的数据之间如何关联、在数据更新过程中如何检验数据的准确度等问题. 对于数据使用, 一方面需要提供快捷高效的搜寻方式供各种需求的使用者便利地获取所需数据; 另一方面需要开发对数据之间有效信息进行挖掘的研究工具, 拓展数据库中数据的应用价值.

#### 4.2 建立多层次电池材料数据库

电池器件的性能不仅与电池材料的本征性质相关, 也与电池材料的微观形貌、多种材料之间的相互作用<sup>[47]</sup>、外界环境场及器件的宏观构造等多种不同空间和时间尺度上的性质紧密关联, 因此要获得从材料性质到器件性能之间的认识, 需要建立多层次的电池材料数据库<sup>[48,49]</sup>. 表 4 列出了在原子尺度、微观尺度、外场效应、多相作用和宏观尺度上所涉及的电池材料数据及可能的用途和使用方法.

## 5 结 论

在材料基因组思想所推动的高通量技术发展下, 电池材料数据库获得了快速的发展, 在计算方法多样性、数据完备性和各类关键性质数据的获取方面都有进展. 未来, 电池材料数据库在提供材料数据的基础上, 将进一步面向应用需求, 构建不同层级的电池数据, 并整合嵌入通用的机器学习算法, 实现研究人员从数据获取、数据挖掘到数据预测的新材料探索过程. 电池材料数据库的建立将有效地提升基于材料基因组的科研数据的有效管理及公共服务能力, 对于与能源材料探索以及与电子、离子输运相关的物理性质的理解都会起到积极

的作用, 同时也将为在材料研究领域引入人工智能方法提供必不可少的数据基础.

#### 参考文献

- [1] Armand M, Tarascon J M 2008 *Nature* **451** 652
- [2] Chen R S, Li Q H, Yu X Q, Chen L Q, Li H 2020 *Chem. Rev.* **120** 6820
- [3] Jain A, Persson K A, Ceder G 2016 *APL Mater.* **4** 053102
- [4] Suh C, Fare C, Warren J A, Pyzer-Knapp E O 2020 *Annu. Rev. Mater. Res.* **50** 1
- [5] Mueller T, Hautier G, Jain A, Ceder G 2011 *Chem. Mater.* **23** 3854
- [6] Kirklin S, Meredig, Wolverton C 2013 *Adv. Energy Mater.* **3** 252
- [7] Xiao R J, Li H, Chen L Q 2015 *Sci. Rep.* **5** 14227
- [8] Rasmussen F A, Thygesen K S 2015 *J. Phys. Chem. C* **119** 13169
- [9] Sikora B J, Wilmer C E, Greenfield M L, Snurr R Q 2011 *Chem. Sci.* **3** 2217
- [10] Ashton M, Paul J, Sinnott S B, Hennig R G 2017 *Phys. Rev. Lett.* **118** 106101
- [11] <http://e01.iphy.ac.cn/bmd> [2020-9-17]
- [12] Zhang L W, He B, Zhao Q, Zou Z Y, Chi S T, Mi P H, Ye A J, Li Y J, Wang D, Avdeev M, Adams S, Shi S Q 2020 *Adv. Funct. Mater.* **30** 2003087
- [13] Korbelt S, Marques M A L, Botti S 2016 *J. Mater. Chem. C* **4** 3157
- [14] Avdeev M, Sale M, Adams S, Rao R P 2012 *Solid State Ionics* **225** 43
- [15] Gaultois M W, Sparks T D, Borg C K H, Seshadri R, Bonificio W D, Clarke D R 2013 *Chem. Mater.* **25** 2911
- [16] Carrete J, Li W, Mingo N 2014 *Phys. Rev. X* **4** 011019
- [17] Deem M W, Pophale R, Cheeseman P A, Earl D J 2009 *J. Phys. Chem. C* **113** 21353
- [18] Zhang T T, Jiang Y, Song Z D, Huang H, He Y Q, Fang Z, Weng H M, Fang C 2019 *Nature* **566** 475
- [19] Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G, Persson K A 2013 *APL Mater.* **1** 011002
- [20] Curtarolo S, Setyawan W, Hart G L W 2012 *Comput. Mater. Sci.* **58** 218
- [21] Saal J E, Kirklin S, Aykol M, Meredig B, Wolverton C 2013 *JOM* **65** 1501
- [22] <https://atomly.net> [2020-9-17]

- [23] Ghadbeigi L, Harada J K, Lettiere B R, Sparks T D 2015 *Energy Environ. Sci.* **8** 1640
- [24] Huang S, Cole J M 2020 *Sci. Data* **7** 260
- [25] Li Y S, Qi Y 2019 *Energy Environ. Sci.* **12** 1286
- [26] Tian H K, Chakraborty A, Talin A, Eisenlohr P, Qi Y 2020 *J. Electrochem. Soc.* **167** 090541
- [27] Peng J Y, Zu C X, Li H 2013 *Energy Storage Sci. Technol.* **2** 55 (in Chinese) [彭佳悦, 祖晨曦, 李泓 2013 储能科学与技术 **2** 55]
- [28] Zu C X, Li H 2011 *Energy Environ. Sci.* **4** 2614
- [29] Wu J Y, Liu P, Hu Y S, Li H 2016 *Energy Storage Sci. Technol.* **5** 443 (in Chinese) [吴娇杨, 刘品, 胡勇胜, 李泓 2016 储能科学与技术 **5** 443]
- [30] Wang L, Wu Z, Zou J, Gao P, Niu X, Li H, Chen L 2019 *Joule* **3** 2086
- [31] Cao W, Zhang J, Li H 2020 *Energy Storage Mater.* **26** 46
- [32] Ceder G 2011 *MRS Bulletin* **35** 693
- [33] Kirklin S, Saal J E, Meredig B, Thompson A, Doak J W, Aykol M, Rühl S, Wolverton C 2015 *NPJ. Comput. Mater.* **1** 15010
- [34] Jain A, Hautier G, Ong S P, Persson K 2016 *J. Mater. Res.* **31** 977
- [35] Hachmann J, Olivares-Amaya R, Atahan-Evrenk S, Amador-Bedolla C, Sánchez-Carrera R S, Gold-Parker A, Vogt L, Brockway A M, Aspuru-Guzik A 2011 *J. Phys. Chem. Lett.* **2** 2241
- [36] Shi S Q, Gao J, Liu Y, Zhao Y, Wu Q, Ju W W, Ouyang C Y, Xiao R J 2016 *Chin. Phys. B* **25** 018212
- [37] He B, Chi S, Ye A J, Mi P H, Zhang L W, Pu B W, Zou Z Y, Ran Y B, Zhao Q, Wang D, Zhang W Q, Zhao J T, Adams S, Avdeev M, Shi S 2020 *Sci. Data* **7** 151
- [38] Nuspl G, Takeuchi T, Weiß A, Kageyama H, Yoshizawa K, Yamabe T 1999 *J. Appl. Phys.* **86** 5484
- [39] He B, Ye A J, Chi S, Mi P H, Ran Y B, Zhang L W, Zou X X, Pu B W, Zhao Q, Zou Z Y, Wang D, Zhang W Q, Zhao J T, Avdeev M, Shi S 2020 *Sci. Data* **7** 153
- [40] Adams S, Rao R P 2011 *Phys. Status Solidi A* **208** 1746
- [41] Henkelman G, Jonsson H 2000 *J. Chem. Phys.* **113** 9978
- [42] Kirklin S, Meredig B, Wolverton C 2013 *Advanced Energy Materials* **3** 252
- [43] Zhu Y, He X, Mo Y 2017 *Adv. Sci. (Weinh)* **4** 1600517
- [44] Wang X L, Xiao R J, Li H, Chen L Q 2016 *Phys. Chem. Chem. Phys.* **18** 21269
- [45] Sendek A D, Cubuk E D, Antoniuk E R, Cheon G, Cui Y, Reed E J 2018 *Chem. Mater.* **31** 342
- [46] Liu B, Yang J, Yang H, Ye C, Mao Y, Wang J, Shi S, Yang J, Zhang W 2019 *J. Mater. Chem. A* **7** 19961
- [47] Wang A P, Kadam S, Li H, Shi S Q 2018 *NPJ Comput. Mater.* **4** 15
- [48] Liu Y, Zhao T L, Ju W W, Shi S Q 2017 *J. Materiomics* **3** 159
- [49] Liu Y, Guo B R, Zou X X, Li Y J, Shi S Q 2020 *Energy Storage Mater.* **31** 434

SPECIAL TOPIC—Fundamental physics problems in all solid state batteries

## Development and application of battery materials database\*

Wu Si-Yuan<sup>1)2)</sup> Wang Yu-Qi<sup>1)2)</sup> Xiao Rui-Juan<sup>1)2)†</sup> Chen Li-Quan<sup>1)2)</sup>

1) (*Key Laboratory For Renewable Energy, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*)

2) (*University of Chinese Academy of Sciences, Beijing 100049, China*)

( Received 17 September 2020; revised manuscript received 21 October 2020 )

### Abstract

High-throughput methods based on automation technology and computer technology can quickly provide tens of thousands of scientific research data, which poses a new challenge to the scientific and efficient management of scientific data. Rechargeable secondary batteries are the keys to the development of electric vehicles and the first choice of wind/photoelectric energy storage. The discovery of new battery materials plays an important role in improving the performance of the secondary batteries. New methods based on big data can be introduced into the screening and design of battery materials to accelerate the development of secondary batteries. This work introduces the development and application of battery material database from the aspects of data acquisition, construction of general and specific battery material database, and the challenges faced by the battery material database.

**Keywords:** battery materials, high-throughput calculations, database, materials genome

**PACS:** 61.68.+n, 66.10.Ed, 82.47.Aa

**DOI:** [10.7498/aps.69.20201542](https://doi.org/10.7498/aps.69.20201542)

---

\* Project supported by the National Key R&D Program of China (Grant No. 2017YFB0701600) and the National Natural Science Foundation of China (Grant No. 51772321).

† Corresponding author. E-mail: [rjxiao@iphy.ac.cn](mailto:rjxiao@iphy.ac.cn)