

电动汽车产业数据库建设初探

柳文灿 杨帆 孔德洋 符钢战

(同济大学汽车学院, 同济大学新能源汽车产业化研究中心, 上海 200092)

摘要: 综述国内外电动汽车产业数据库建设现状, 指出我国电动汽车产业数据库建设上的不足, 研究电动汽车产业数据库建设的需求和数据源问题。基于电动汽车产业的系统性和数据库的建设流程, 从车辆和应用环境两个维度构建了电动汽车产业数据库内容的“VP+ISBCG”框架, 最后对电动汽车数据库的功能和应用进行分析。全文研究了电动汽车产业数据库建设中所涉及的问题及应用前景, 为电动汽车产业数据库平台的建设提供了科学的思路和系统理论。

关键词: 电动汽车; 产业数据库; 应用环境; 数据库建设; 信息资源

中图分类号: G304

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2014.04.005

Primary Discussing to Construction of Electric Vehicles Industry Database

Liu Wencan, Yang Fan, Kong Deyang, Fu Gangzhan

(School of Automotive Studies, Centre for Automotive Industry(CAI), Tongji University, Shanghai 200092)

Abstract: By reviewing the present situation of domestic and foreign electric vehicle industry databases, this paper identifies shortcomings of China's electric car industry database construction and analyzes the needs of the construction of the electric vehicle industry databases and data sources. Based on the construction process of database and systematic character of electric vehicle industry, the electric car industry database contents framework "VP + ISBCG" has been built from the two angles: the vehicle and the application environments. Finally, it discusses the functions and application problems of electric vehicles database. This paper shows a relatively comprehensive study of the problems and prospects of electric car industry involved in building the database, providing scientific ideas and systems theory for the construction of the electric car industry database platform.

Keywords: electric vehicle, industry database, application environment, database construction, information resources

1 引言

我国电动汽车随着示范运行工作的展开进入了新阶段, 产业规模不断扩大, 在电动汽车技

术研发和产业化推进活动中, 众多的信息资源及产业数据正不断形成和积累, 这些分散在各类机构和产业链不同环节的产业数据, 涉及范围广、信息量大、形式多样, 搜集难度较大, 需要借助

作者简介: 柳文灿 (1990-), 女, 同济大学硕士研究生, 研究方向: 新能源汽车市场与商业模式; 杨帆 (1987-), 男, 研究助理, 研究方向: 车联网技术与汽车共享; 孔德洋* (1973-), 男, 同济大学副教授, 研究方向: 汽车市场信息与新能源汽车产业; 符钢战 (1958-), 男, 同济大学教授, 研究方向: 新能源汽车产业。

基金项目: 国家科技支撑计划项目“电动汽车技术预测与决策支持系统开发”(2013BAG06B00); 国家国际科技合作项目“基于城市工况运行的电动汽车关键技术国际合作研究”(2011DFB80100); 同济大学“985”重点建设项目“新能源汽车产业数据库建设与应用研究”(119220)。

收稿日期: 2014年5月19日。

信息技术加强电动汽车相关信息资源的整合与共享,并对信息进行分类、加工、整理、提炼,以及及时反映电动汽车领域的国际动态和前沿技术。因此,为有效整合和集成我国产业信息资源,十分有必要建立电动汽车产业数据库。

本文在分析国内外电动汽车产业数据库建设现状的基础上,重点研究电动汽车产业数据库的需求、数据源、产业数据库建设内容和结构以及数据库的应用,为电动汽车产业信息化建设提供参考。

2 国内外电动汽车数据库简况

日本建立了专门的研究机构,如日本新能源·产业技术综合开发机构(NEDO),投入大量资金进行电动汽车数据库和信息决策系统建设,为政府根据每年技术和产业发展动向而调整下一年电动汽车发展战略和动力电池技术路线图提供了巨大的支持。美国建立了规范的各种车型的(包括混合动力、电动汽车)燃油经济性数据库。该数据库包含车辆原理原本数据和工况测试数据,可对任何车型进行检索以及性能比较分析。

国际上主要的电动汽车数据库有:SAE数据库是美国机动工程师协会开发的数据库,涵盖有全球汽车主题索引数据库(GMD)、SAE技术论文(16000篇)、SAE标准全集;全球混合动力及电动汽车数据库能够提供模式智能的市场工程化预测,同时该数据库也能够提供OEM级别的信息,也可以根据区域、国别、OE类型、汽车类型和部件类型进行浏览查询;IHS公司的混合动力及电动汽车数据库提供全方位的混合动力及电动汽车产业数据,能够为会员在快速变化的细分市场中提供最新上市的汽车分析资讯;EV-INFO公司的EV公司数据库包括了世界各地9000多家电动汽车公司,涉及研发、制造和销售等环节。

我国电动汽车技术经过10余年的建设,初步建立了电动汽车技术体系,已申请专利3000余项,颁布电动汽车国家和行业标准56项,建成30多个节能与新能源汽车技术创新平台。目前,有关电动汽车的信息主要分布在两类系统:一部

分分布在国内的各类电动汽车网站,如第一电动网建立了新能源汽车数据库,为电动汽车相关企业提供企业信息、产品信息、交易信息;另一部分则分布在国内相关研究机构和信息资源机构,如广东省知识产权公共信息综合服务平台集成了电动汽车专利数据库,能够从电动汽车模块、企业、充电设施等进行专利检索。

同发达国家相比,一方面,我国在数据采集上比较单一、不全面。例如在示范运行数据采集上,国内2008年北京奥运和2010年上海世博会期间开展了大规模电动汽车示范运行,但由于受时间限制和示范车辆仅为国内车辆,采集的数据统计长度不够,也缺乏与国际同类车辆的比较;另一方面,我国尚缺乏综合、系统的电动汽车数据库,无论是数据量、数据来源、数据应用、数据规范等方面,同美、日等国家相比都有不小的差距。因此,建立系统的、多源信息综合利用的电动汽车数据库,为电动汽车产业发展提供有效的信息服务和决策参考依据,成为电动汽车信息基础设施建设的重要内容。

3 产业数据库的建设需求与数据源分析

数据库是以一定的数据模型组织和长期存储在计算机内的大量共享数据的集合,是计算机应用系统中一种专门管理数据资源的系统^[1]。建设电动汽车数据库要结合电动车领域的实际特点,同时也要遵循一定的建设和设计流程,以保证数据库的优良性。在建设前期根据数据库特色进行需求分析和数据源分析非常重要。

(1)产业数据库需求:在建立产业数据库前的首要任务是分析产业界的相关需求,其主要体现在两个方面,一是信息服务的需求。电动汽车信息资源的膨胀使得获取特定信息的难度加大,而电动汽车企业、用户或科研人员对特定的产品、市场和技术信息需求不断上升,因此,围绕产业领域对信息资源进行专题收集与加工,开发出特色专题数据库,形成定向的信息产品,能更好地满足企业、消费者或科研人员与日俱增的信息需求。二是决策支持需求。电动汽车相应的技

术研发与产业化模式同传统汽车产业差异较大，电动汽车的管理决策需依托以事实型数据资源为基础的科学战略平台，对多源信息进行集成整合，借助数据挖掘工具和方法以及专家团队智慧，构筑全景决策环境，从而最大限度地促进电动汽车领域管理决策的科学化，提高决策效率。

(2)数据源：电动汽车产业的数据源来自产业发展中的各个环节，各类数据按内容主要划分为两类，一是科研学术数据，一是推广应用数据。其中，科研学术数据来源为国内外学术期刊、论文、专利、技术报告等，涵盖电动汽车技术发展路径及技术预测模型；推广应用数据来源为电动汽车产品信息、电动汽车车联网信息、电动车技术标准信息、电动汽车计划项目信息、电动汽车相关机构信息、电动汽车人才信息以及电动汽车政策法规信息等。电动汽车产业信息资源之间具有极强的相关性，具有较高的集成分析和利用价值，但是数据形式和来源途径的多样化使得数据收集、处理和分析存在一定的难度，其数据库建设结构的顶层设计对其数据库体系的建立至关重要。

此外，在数据组织和管理上，遵循一定的建库原则，如：数据的收集一般以时间坐标轴集合；以面向产业链的各环节的关键活动来组织信息；强调数据的透明度与可靠性；强调数据收录的标准化、国际化以及扩展的可能。随着数据库技术在实践工作中的广泛应用，出现了许多面向专门应用领域的数据库技术，如工程数据库、统计数据库、科学数据库、空间数据库、地理数据库等，这类数据库在原理上也没有多大的变化，

但是它们却与一定的应用相结合^[2]。针对电动汽车领域，数据库的建设通过对产业数据源的分析，需着眼于数据库基本结构的构建，以反映应用需求，方便与用户进行交流。

4 数据库内容和结构

电动汽车产业数据涉及电动汽车整条产业链的各个环节，包括电动汽车技术研发、产品发展、充换电基础设施、维保服务体系等多个环节，因此，本文在相关项目研究和电动汽车产业链研究的基础上对电动汽车产业进行模块划分，通过对每个模块数据库构成的分析，完成电动汽车产业数据库建设的结构设计，建设结构如图1所示。

电动汽车数据可以分为两大类：一类数据是与电动汽车本身的技术和产品相关，包括整车和零部件数据。另一类数据是与电动汽车应用环境相关，包括基础设施、服务体系、商业模式、用户认知和政策法规等数据。电动汽车产业的发展不仅仅依赖技术进步和产品完善，应用环境体系的建设和完善也非常重要，二者相互促进，相辅相成。因此在电动汽车产业数据库建设过程中，也必须包含这两大类的数据信息。综合起来看，就是“VP+ISBCG”数据结构，其中，车辆数据包括电动汽车车型数据包(V)和关键零部件数据包(P)2个子模块，应用环境数据包括充换电基础设施数据包(I)、电动汽车维保服务数据包(S)、电动汽车商业化数据包(B)、电动汽车用户认知数据包(C)和政策法规数据包(G)共5个子模块。

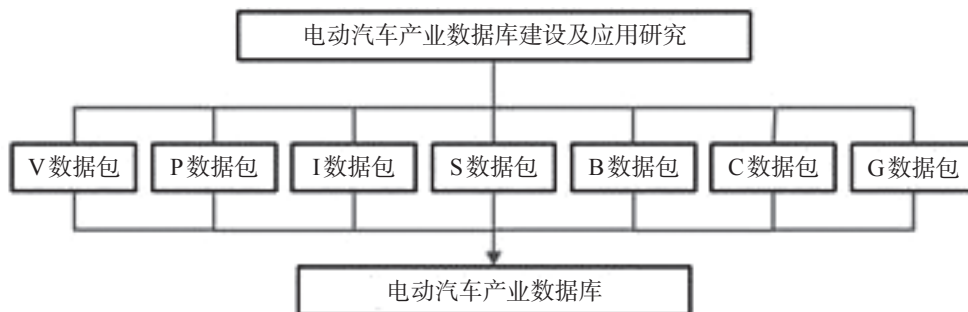


图1 电动汽车产业数据库建设结构

(1) 车型数据包 (V)。该数据包包括国内外主要汽车企业的混合动力、纯电动、插电式混合动力、增程式电动车、燃料电池汽车的车型信息(整备质量、车身参数等),包括产品的性能参数(动力系统额定功率、储能系统额定电压、标称容量等)、技术方案(电池、电机等关键零部件的种类、配套厂家、技术参数)、产品状态(上市时间、地区)、市场信息(价格、产销量)。通过国内电动汽车展览会、汽车期刊、年鉴、行业报告、其他车型数据库及网站等资源,实现车型信息形式的格式化和数据化,从而方便查询和动态更新。

(2) 关键零部件数据包 (P)。该数据包主要搜集和整理国内外主要的三电(电池、电机、电控)关键零部件厂商和产品信息,包括主要生产厂商的基本情况(名称、性质、注册资金、上市情况、生产基地等)、产品的性能参数(型号、额定电压、技术方案等)、产品状态(上市时间、地区等)、市场信息(价格、产销量)。通过对电动汽车关键零部件产业数据库的建设,了解关键技术发展情况以及国内外研究动向,可提供有效的行业情报和产业指导政策的相关依据。

(3) 充换电基础设施数据包 (I)。该数据包主要包括电力、电池、充换电设施、充换电服务运营技术数据、供应和市场及供应商数据,数据内容涉及充换电站设施技术参数和关键设备、服务运营投资主体、运营主体、运营模式等。电动汽车的结构和核心技术与传统燃油汽车有本质上的不同,其产业链系统的复杂程度远超过传统燃油车。该数据包的建设首先需要确定采集对象范围及内容并确定采集方式,对采集数据进行整理和加工,按照指定的规范对数据进行加工处理,同时要定期对配套基础设施及使用状况数据库进行更新。

(4) 维保服务数据包 (S)。该数据包包括各地维保服务的案例,多种不同服务商的维保体系、维保标准及规范等,涉及电动车车联网数据库的建设。通过“人-车-路”信息的互通,为电动车用户提供及时可靠的电子信息平台服务体

系。作为未来智能交通的重要组成部分,车联网是以车为节点的信息系统,综合现有的电子信息技术,将每一辆汽车作为一个信息源,通过无线通信手段连接到网络中,进而实现对全国范围内车辆的统一管理^[3]。这也使得IT服务供应商、内容提供商、车联网设备提供商扮演着重要角色,越来越多的互联网企业开始进入汽车行业的产业链^[4]。数据包的主要难点在于建立和整理出一个通用的数据结构。

(5) 商业化数据包 (B)。该数据包涵盖国内外电动汽车量产车型商业化数据,包括商业化推广模式、各利益相关方的合作关系、推广的规模等。燃油汽车的设计和改型,可以由汽车制造企业独立完成,制造商通过与用户的互动,就可以获得燃油汽车发展所需要的信息和资源。而电动汽车的商业化,却要求在汽车制造商、电池制造商、充电运营商以及用户的互动中才能找到发展的方向和技术解决方案^[5]。通过对有效的电动汽车商业模式相关数据的收集,对分析电动汽车技术成熟度、规划电动汽车发展技术路线起到重要支撑作用。

(6) 用户认知数据包 (C)。该数据包针对用户使用和认知情况,构建用户的基本信息,包括年龄、性别、收入、车辆拥有情况等分类,从而考察不同人群其对不同电动汽车车型的驾驶感受、外观内饰、操控性能等的感受,进而推测其购买意愿。该数据包可通过实地及网络的调查问卷取得。数据包的难点在于形成稳定形式的调查问卷,以此通过比较多次的调查结果来反应用户对电动汽车认知度的变化。

(7) 政策法规数据包 (G)。该数据包主要用来分析国内外、国家及地方层面的电动汽车直接和间接扶持、补贴、税惠等政策,包括法规名称、法规文号、法规关键词、法规全文等内容、颁布机构、生效时间、效力等内容,同时还收录国内外电动汽车领域已出台的技术标准。该数据包主要通过人工采集的方式,从不同的网站分别收集相关文件,以人工的方式进行解读和归类,并抽取相关的内容构建数据包。数据包的难点在

于政策法规文件大量分散于各级管理机构，面临文件的归类和去重等问题。

在建设数据库时，对于能够从已建成的数据库中抽取获得的信息，首先确定检索策略，然后通过全面检索各类相关主题数据库，如世界各国、地区和国际知识产权组织公布的电动汽车相关的专利申请，随后结合电动汽车技术领域范围，利用专家智慧和工具软件对相关数据进行分类和加工，形成数据库。

对于网络信息资源，特别是动态信息采集方面，主要通过为不同的数据源定制抓取代理；为每个数据源注册监听器，从而做到快速定位新的数据资源，并更新电动汽车行业数据库。对于未形成数据库的数据，通过数据合作、数据购买、联合开发及人工调研等多种方式，获取数据，建立数据库。

5 数据库的应用

构建电动汽车产业数据库有助于将电动汽车产业发展过程中的各项活动、项目及研究成果的多源信息和数据资源进行有效集成和整合，实现电动汽车产业体系信息化建设的整体性和连续性，避免系统孤岛、信息碎片的形成^[6]，同时借助挖掘工具和方法，将数据有效应用于各个领域，发挥数据的潜在功能。其具体可以实现如下几方面的应用(图2)。

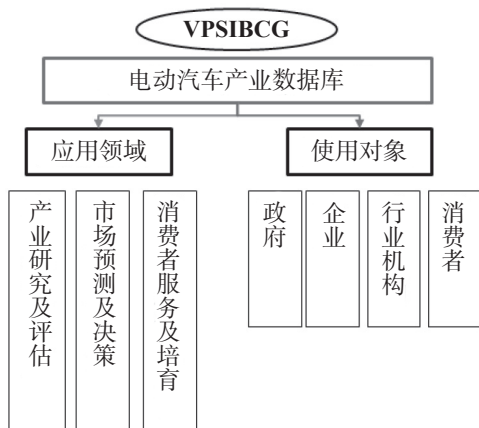


图2 电动汽车产业数据库的应用

(1) 产业研究及评估。行业研究机构或政

府、企业可基于数据库的技术和应用环境两大类数据，对电动汽车关键技术进行验证，对电动汽车行业技术动态及技术演变路径进行跟踪，以促进关键技术和相关产品的研发及投产制造；在应用环境方面通过对电动汽车产业链上零部件、整车及运营企业产品、市场推广、运营状况等信息进行整合分析，发现应用推广环境的薄弱环节从而创新商业模式及运营机制。同时，产业数据是进行科学评估的基础，如通过对示范城市推广情况、技术设施建设情况、产业带动情况等数据包的有效建立，可以客观评估城市的示范运行系统和示范运行效果，为科研项目的验收提供事实依据。

(2) 市场预测及决策。国家、政府和行业研究部门在规划电动汽车产业、制定相关政策、预测技术走向及进行相关决策的过程中，需借助产业数据库以达到数据、工具与专家协同决策的管理目标，建立基于“事实数据+工具方法+专家智慧”的综合决策模式，使管理决策者和专家团队能够客观了解行业发展现状，充分利用数据工具，同时也最大限度地发挥专家群体的智慧，提升决策的科学化水平和效率。

(3) 消费者服务及培育。通过形成权威的电动汽车相关企业及运营单位绩效排名、产品排名、品牌指数等可以让消费者更直观地了解电动汽车市场，提供产品信息、市场动态等高质量、有序化的信息服务，逐步培育潜在用户和开发细分市场。

因此，对于国家和地方而言，建设电动汽车产业数据库，对发展本土电动汽车产业，具有重要的基础性支撑作用，是我国在电动汽车领域参与国际竞争、开展国际合作、进行政府决策及地方规划的基础。对于汽车行业和企业而言，基于数据平台，开展指数、预警及评价体系等研究，对加快产业结构调整，助推企业实现低碳转型、提高核心竞争力，具有积极的引导作用。随着电动汽车产业发展进程的不断加快，电动汽车产业数据库在未来将会有着广阔的应用前景。

6 结语

电动汽车产业数据库的建设在国际上仍然处于起步阶段,随着电动汽车产品应用的普及和产业发展的深化,建设电动汽车产业数据库,实现车辆信息资源和应用环境信息资源的数据化已成为一项重要的基础性工作。本文基于电动汽车产业链的发展特点,从技术和应用环境两大角度系统地构建了“VP+ISBCG”产业数据库框架,填补了我国电动汽车产业信息资源建设领域的空白。

在大数据时代背景下,随着数据分析与数据挖掘技术的进一步发展,产业数据库可以为不同领域的决策者或者用户如电动汽车用户、科研机构、政府等提供权威的信息服务,满足及时的、定制化的数据需求,既必要又可行,对我国目前的发展阶段而言,这既是挑战又是机遇,在参照

国外同行数据库建设经验基础上,需要加强研究,以建设符合我国自身发展特色的电动汽车产业数据库。

参考文献

- [1] 李跃明.数据库系统的创新发展[J].电脑知识与技术,2011,7(2):269-270,273.
- [2] 向海华.数据库技术发展综述[J].现代情报,2003(12):31-33.
- [3] 叶瑞克,陈秀妙,朱方思宇,等.“电动汽车-车联网”商业模式研究[J].北京理工大学学报:社会科学版,2012,14(6):39-44.
- [4] 许海洋.电动汽车产业链格局探究[J].汽车工业研究,2013(11):22-26.
- [5] 隋忠海,王震坡.对我国电动汽车产业链发展模式的思考[J].企业研究,2011(4):58-61.
- [6] 郭理桥.建设行业多层次数据库建设的思考[J].中国建设信息,2009(22):6-9.

(上接第24页)

针对电动汽车多源信息决策支持系统完善的数据质量管理流程,对系统内的数据质量实施全过程、全领域管理,将数据质量管理以制度化、规范化的方式落实到数据生成、传递和使用的各个过程之中。相对以往的数据质量管理流程,它细化了管理流程,更加注重对研究人员数据质量意识的培养,要求提高系统内各数据库建设人员的协同配合。按此流程实施后,在一定程度上做到了数据质量从源头开始控制,优化了数据字段设置,提高了数据采集加工入库的速度,系统内数据能够做到可检索、可统计、可展示,为电动汽车多源信息决策支持系统提供有力的数据支撑保障。

参考文献

- [1] 商广娟.有效的数据质量管理体系——21世纪管理

- 的基石[J].航空标准化与质量,2005(2):18-22.
- [2] 常宁.IMF的数据质量评估框架及启示[J].统计研究,2004(1):27-30.
- [3] 侯瑜.基于DQAF框架的我国统计数据质量管理及改进[J].统计研究,2012(12):24-30.
- [4] 宋敏,覃正.国外数据质量管理研究综述[J].情报杂志,2007(2):7-9.
- [5] 韩京宇,徐立臻,董逸生.数据质量研究综述[J].计算机科学,2008(2):1-5,12.
- [6] 郭志懋,周傲英.数据质量和数据清洗研究综述[J].软件学报,2002(11):2076-2082.
- [7] 叶宏伟,金中仁,陈振宇.图书馆信息集群研究[J].中国图书馆学报,2008(1):99-101.
- [8] DAMA International. DAMA数据管理知识体系指南[M].北京:清华大学出版社,2012.
- [9] 孙中东.企业级数据治理框架下的数据质量管理[J].金融电子化,2011(6):57-60,6.
- [10] 谷斌.信息系统建设中的数据质量管理体系研究[J].情报杂志,2007(5):65-67.