DOI:10.13216/j.cnki.upcjess.2020.03.0001

石油工程大数据应用的挑战与发展

苏　健[1,2],刘　合[1]

(1. 中国石油勘探开发研究院,北京 100083; 2. 中国石油大学(华东) 经济管理学院,山东 青岛 266580)

摘要:近年来,随着国内油气勘探开发工作的不断深入,我国油气行业在地质、测井、物探、开发等各个环节均积累了海量的油田数据资源,油田的信息化、数字化、智能化建设也逐步展开,油气勘探开发已经进入"大数据"时代。大数据作为多学科一体化"融合"的纽带,对石油工程具有重要的意义,但因其数据量巨大、结构复杂、形式多样、种类繁多等特点,目前没有形成统一的、及时更新的、共享的油气井地质与工程基础数据库,并且缺乏对数据的分类整理和二次开发。面对当前大数据在石油工程应用及一体化建设中的挑战,提出合理的数据处理方式及发展思路,对于石油工程大数据更好地应用于油田生产实践,提高油田生产决策的科学性具有一定的指导意义。

关键词:大数据;石油工程;关键技术;管理模式

中图分类号:TE319; TP311.13　**文献标识码:**A　**文章编号:**1673-5595(2020)03-0001-06

一、引言

大数据指的是需要新处理模式才能具有更强的决策力、洞察力和流程优化能力的海量、高增长率和多样化的信息资产,具有大体量(Volume)、时效佳(Velocity)、多样性(Variety)、价值密度低(Value)、真实性(Veracity)的 5V 特点。[1]"大数据"在 2010 年前后就已经受到许多国家及行业的重视,例如美国政府在 2012 年宣布启动"大数据研究和发展计划",以此增强收集数据、分析挖掘信息的能力,希望带动相关产业发展,并将"大数据战略"上升到了国家战略的层面[2],这也预示着"大数据"时代的到来。

随着勘探开发的不断深入,我国油气行业在地质、测井、物探、开发等各个环节均积累了海量的数据资源,随着油田的信息化、数字化、智能化建设,对油气生产过程中积累的数据资源的分析利用也逐步展开,油气勘探开发步入"大数据"时代。石油工程大数据的挖掘,是对油气生产过程中各个环节的数据进行充分的信息采集,并且对所采集到的信息进行全方位的分析利用,进而可以对生产环节中指标的变化趋势、工况效率、工艺流程的安全性以及成本控制等方面提出预测、预警和优化,运用已获取的信息、规律、趋势,适时调整生产方案和策略,达到提升油气产量和效益的目的。[3]

由于前期积累的数据量巨大、结构复杂、形式多样、种类繁多,并且数据不连贯,没有形成统一的、及时更新的、共享的油气井地质与工程基础数据库,对数据缺乏分类整理和二次开发,形成了"数据孤岛"。因此,大数据在石油工程及地质工程一体化等方面仍存在许多挑战,需要通过更科学的管理手段、数据结构优化和体制机制建立来推进其发展进程。

二、石油工程大数据应用的意义及面临的挑战

(一)大数据应用的重要作用

大数据分析是一个对数据进行收集、清理、转换和建模以发现有用信息的过程,这些信息可用于为未来的决策提供建议。大数据分析正在从根本上改变某些行业,如零售营销、电信、保险和银行。在数

收稿日期: 2020-03-06

作者简介: 苏　健(1988—),男,山东东营人,中国石油勘探开发研究院工程师,中国石油大学(华东)经济管理学院博士研究生,研究方向为能源战略规划。

通信作者: 刘　合(1961—),男,黑龙江哈尔滨人,中国石油勘探开发研究院教授级高级工程师,博士,博士生导师,863 计划首席科学家,中国工程院院士,研究方向为油气田开发、能源与矿业工程管理。

字时代,企业利用技术来改变运营方式变得越来越重要,这得益于大数据分析。近年来,数字技术在油气勘探生产中的应用也越来越多。

大数据作为多学科一体化融合的载体和纽带,贯穿石油工程的全过程,包括勘探开发过程中从钻井到生产等一系列工程技术应用及解决方案筛选优化等工作,同时贯穿于油气藏预探、评价、开发、调整的全生命周期,对石油工程以及地质工程一体化建设与应用均具有重要意义。

(1)大数据可应用于地质工程多学科数据梳理及研究,紧紧围绕与地质工程相关的参数及数据体,直接服务于提产、增效的具体环节,并且在此过程中,相关认识、数据体可在生产中应用、验证、再优化。

(2)大数据可应用于实用性工艺技术的合理化优选,通过加强现场技术支持、加快数据分析及数据体的更新,可形成动态决策机制,将具有“时效性、真实性”的数据及认识,渗入到生产施工的优化中。

(3)通过深入的“数据挖潜”和高效的“数据应用”,可以快速、动态得到结果,并不断验证,从而实现效果持续改良与提升,实现一体化施工推进的提产与增效。

(二)大数据应用面临的挑战

大数据分析的快速发展为公司的发展提供了机会,为了降低成本、提高效率和安全性,许多油气企业正在逐步加大努力,抓住这些机会,将大数据转换为有用的信息并及时做出决策。但随着数据量和数据多样性的持续增长,如何有效地利用数据仍需要认真考虑,数据的利用必须基于数据是透明和可信任的,但如何确保数据的准确性仍然是很大的挑战。

油气田企业的动静态数据是海量的。以石油上游统建系统为例,A1 到 An 条块分割式数据管理模式积累了海量的数据资源,涵盖了勘探与生产技术数据管理系统、油气水井生产数据管理系统、采油与地面工程运行管理系统、工程技术生产运行管理系统、勘探与生产调度指挥系统、油气生产物联网系统、工程技术物联网系统等多个系统业务功能,构建了 9 大类数据资产库,覆盖勘探评价开发井近 45 万口,管理测井数据体文件 108 万余个、油气藏 600 个、地震数据 7 000 余个工区、成果文档 367 余万份等,累计存储数据 1.7PB 以上。油气田企业存储了海量的油田数据,但“数据大”并不等同于“大数据”。

问题一,条块分割的管理模式造成了大量的“数据孤岛”。[4] 目前的油气田企业,包含多个条块分割的部门或系统,每个部门或系统都有各自的数据,各部门之间的数据往往都各自存储,自成体系,每个部门或系统的数据就像一个个孤岛一样无法(或者极其困难)和企业内部的其他数据进行连接互动。同一研究的基础数据可能存在于多个系统、多个数据库以及各研究人员的计算机中,无法保证各项研究中基础数据的权威性,影响了研究成果的准确性,不能有效地指导勘探开发研究和生产。

此类数据存在的核心问题是数据库多、平台多、孤立应用多,“数据孤岛”现象严重。[5] 一是数据标准不统一,模式不匹配,没有形成统一的、及时更新的、共享的油气井地质与工程基础数据库,易造成数据冲突、数据冗余,也没有行业内具有兼容性和中立性的专门的咨询服务公司对其进行分类整理和二次开发,而后应用于数据共享、挖掘和分析;二是技术平台的规范不一致,组件不复用,适应性不足,集成共享难;三是系统应用数量多,管理型为主,应用独立,一体化应用少;四是系统建设方面,业务需求响应慢,建设周期长,部分功能重复,系统维护成本高。

“数据孤岛”的直接负面作用包括两个方面:一是投资浪费。后端工程施工结果不能及时反馈给地质部门,油藏模型得不到及时修正,会造成低级错误不断重复,形成极大的浪费。二是效率降低。地质认识的多解性还造成不同专业、不同单位针对同一油气藏工作时,选择使用参数值不一致,研究结果缺乏统一性、可比性;还会经常重复简单低级的资料收集、数据甄别和认识消化工作,严重影响工作效率。

深入分析“数据孤岛”形成的根本原因,主要包括以下几个方面:一是需求不到位。企业在信息化建设时缺乏对员工信息需求的调查了解,把握不够准确,同时,员工自身信息需求的意识较为薄弱。二是标准不统一。许多企业在实施信息化的过程中,根据业务需要开发了满足不同业务需求的多个管理系统,每个系统均单独建立数据库,而不同时期、不同系统的数据库在建立时可能信息标准并不统一,多个数据库之间往往自成体系,并不存在任何联系。三是管理体制问题。由于各业务主管部门之间的差异,并且各部门对于信息化的需求不同,有碍于信息的畅通,原本应该统一的信息数据未能反映到全部的信息系统上。[6]

问题二,数据质量。企业的数据质量与大数据应用之间存在着直接联系,高质量的数据可以使公司保持竞争力并在经济动荡时期立于不败之地。我

们由于 60 年来历史数据收集手段、统计方法、管控手段的不一致性,导致数据的真实性和完整性等都值得磋商,特别是目前上游涉及传统生产测井方法和自动化数据的采集程度还不高,影响了大数据的有效性和精度。

所以说,“数据孤岛”和数据质量制约了工程一体化的深入应用。我们当前数据源的开放程度、加工能力以及数据源的交换水平,对于一体化而言是远远不够的,所以说,目前我们的大数据分析,可能还仅仅停留在概念的层面。

油气藏海量的数据是宝贵的资源,但数据结构复杂、形式多样、种类繁多,缺乏便捷高效和功能型的一体化工作平台与工具。因此,打造一个可以用技术手段穿透部门边界、甄选核心数据的平台,是解决当前存在的“数据孤岛”问题,并使长期积累的海量数据得以高效利用的很好选择。

三、石油工程大数据发展的思路

要成功地应用大数据分析,必须解决各种数据管理问题,克服数据质量的异常值、重复、过时和非结构化等问题。此外,需要将多个不同数据源集成到具有上下文信息的同一平台中。许多公司已经通过构建数据湖、建立数据标准以及数据管理体系来实现这一转变。从根本上说,越来越多的企业正在认识到数据的价值潜力,关注需要收集哪些数据,以及如何收集、处理和存储当前和未来的应用程序的数据。

但是,数据并不在于“大”,而在于“有用”。大数据发展的思路,必须要决策层建立一体化的理念,强化各级管理层的数据化转型认识、知识普及、顶层设计等工作,使各级管理层明确应该如何统筹资源,建立一个便捷高效、路线明晰的数据一体化工作平台,更关键的是要推动平台的有效实施。

(一)构建数据一体化平台

大数据在石油工程领域的发展,应当打破“数据孤岛”,促进一体化数据融合。首先要充分利用既有的数据资源,以油气藏为核心布局各项工作,实现各项参数交互优化、行业间相互推进,要为不同专业和部门沟通搭桥、疏堵破阻,并且记录过程、积累经验、传承知识,才能达到事半功倍的效果。[7]

数据一体化平台的建设应以油田勘探开发油藏评价的业务为对象,依托现有的数据库、专业软件和展示平台,通过全生命周期精细化管理与考评,实现各种数据的图形化展示与对比分析,有效支撑项目研究中地震、地质、开发等多专业的协同研究(一体化平台工作模式见图 1)。

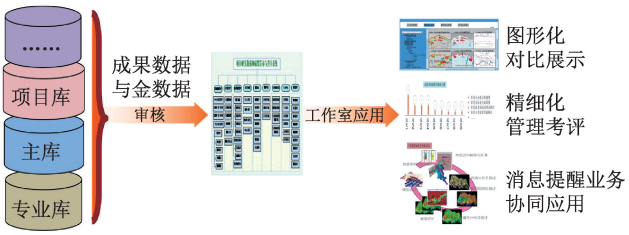

图 1　数据一体化平台工作模式

(二)加强数据质量管控

同时,面对数据质量不断恶化的情况,应当加强对数据质量的管控。在数据采集时,加强“源头治理”,利用物联网技术实现数据的自动化、标准化采集;数据入库时,应重视“数据标准规范”建设,提高数据库的规范性、可用性,减少“手动录入”的入库方式,实现数据库与传感器的标准化对接;数据流通时,应利用区块链技术加强数据流通环节的数据质量管控,确保数据的可追溯、不可篡改;采取数据的分布式存储,实现数据共享,打破“数据孤岛”,促进一体化数据融合。油气勘探开发生产过程理想状态应该是以油气藏为中心,各专业各部门协同配合,用最短的时间、最少的费用掌握地下油气藏的真实赋存状态,采取经济适用的工艺技术定制最合理的装备与工具,实现油气资源勘探开发效益最大化(一体化工作流程见图 2)。

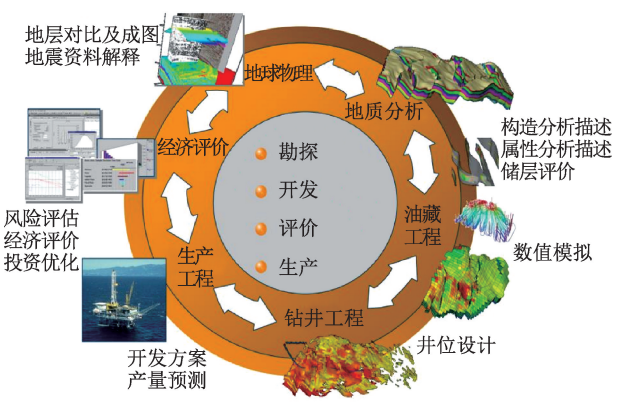

图 2　数据一体化工作流程

(三)建立“油气藏金数据”

应当建立“油气藏金数据”的概念,筛分优选数据资源。“油气藏金数据”主要包含两层含义:一是数据项被不同专业使用的频率高;二是其参数值经过不断优化,数据的准确度高。在油气勘探开发数据系统中,金数据是在整个油藏勘探开发过程中贯穿使用的核心关键数据,对于油气勘探和生产有着重要的作用。通过对油气藏地质数据和工程数据的分析与挖掘,不断总结和更新产能控制因素及改造优化参数,不断改进作业参数和作业流程,从而不断提高钻完井工程效率及储层改造的有效性,持续实

现降本增效。工程实践的突破和再认识,又能够极大推动基础理论研究的深入和发展,使理论与实践的结合和互动得到有效保障。最终通过油气藏金数据实现对现有资源的充分利用。结合企业数字化建设进程,以现有数据库建设为基础,以一体化示范项目为载体,逐步建成大数据系统和一体化支持及协同指挥环境。

对于金数据的寻找,可以分为 5 个步骤(见图 3)。数据资源筛分优选基于业务流程,将研究人员手动管理的研究数据和研究成果以及现有的项目库、专业库、主库等数据库资源进行对比,确定各库数据的一致性,并按照业务模型进行查漏补缺;在此基础上明确金数据范围,分析金数据所属业务等信息;从相关业务表中筛选出金数据字段项;集中针对金数据开展工作,从而达到事半功倍的效果。

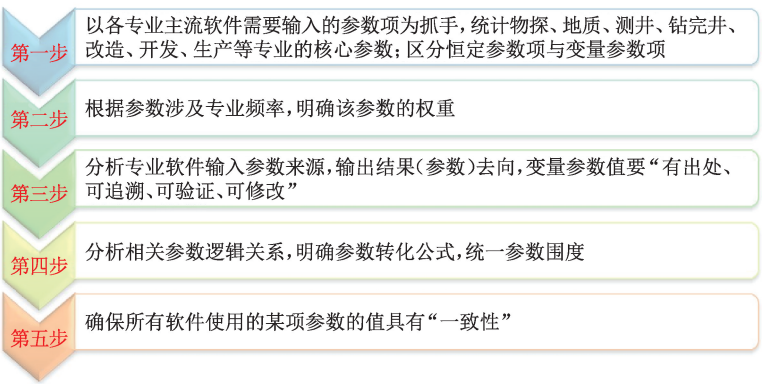

图 3　油气藏金数据筛选步骤

促使地质、油藏、钻完井和地面工程达到一体化融合(见图 4),油藏工程应提前介入,向后延伸到油藏和钻完井工程及地面工程,将钻井、完井、压裂改造、试油、采油、地面集输等工艺一体化融合,钻完井要向前延伸到油藏工程,不断修正地质模型,进而助推新技术工业化、应用规模化发展(见图 5)。

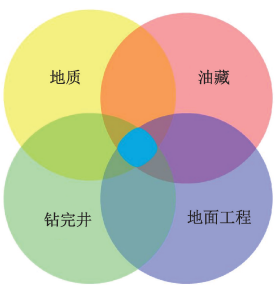

图 4　一体化融合示意图

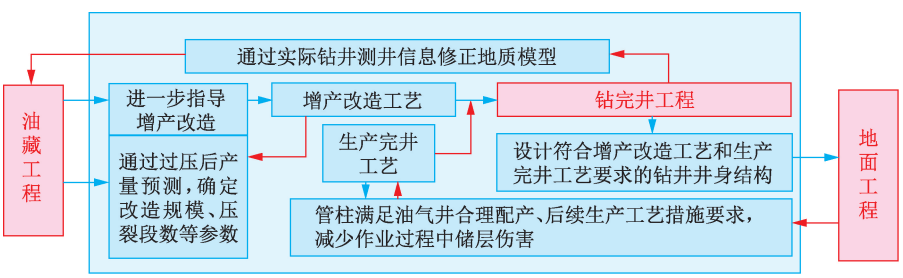

图 5　一体化设计流程

四、几点思考与建议

数十年来,数据作为行业的福音,专家们会想出各种新的数据类型来促使行业发展变得更高效、更安全。每年工业界投入数万亿美元来创建和分析数据,但只有很小一部分投资用于确保这些数据被正确地用于工业。

石油行业是周期性的,随着石油价格的波动,行业的重点往往会转移。[8] 经济下行的挑战导致运营成本被削减,被认为不重要的数据分析功能被外包、缩减甚至取消。通常,数据分析被视为非核心功能,有经验的数据专业人员被解雇,关键的数据存储被下放到专有软件操作中管理。虽然这可能是一种有效的短期成本削减机制,但长期后果往往是毁灭性的。

目前,整个行业正在意识到,如果不首先对数据进行整理,我们将无法在数字化战略举措方面取得任何进展。这需要可靠的数据分析与管理,而不仅仅是靠一点点的数据转换来支持,典型项目的应用要具有普适性。数据在过去常常被视为副产品,而现在被视为关键的业务推动者(甚至是资产)。生

产数据在长时间的相对休眠之后,大数据分析又使它们回到了聚光灯下,越来越多的人感受到数据质量问题的严重性和数据不完整性带来的苦恼,需要有一个可靠的数据管理框架才可以解决这些问题。

针对当前大数据应用面临的挑战及问题,研究提出以下几点建议。

(一)大力发展行业大数据应用,创新管理模式,提升核心竞争力

当前,数字化转型、网络化重构以及智能化提升均在我国不同行业、多个领域积极推进,行业大数据的应用蕴含了巨大的发展潜力和价值,必然是推进企业数字化建设的重要途径和基础。有研究报告指出,许多制造企业在利用大数据技术后,其生产成本能降低 10% ~15%[9],而成本仅仅是大数据技术应用对企业影响的其中一个方面。企业可以通过收集、利用全生命周期中各个环节的数据,更加精细、具体地了解多方需求,建立更加精准、智能的应用系统,改变以往的惯性思维,调整结构,破除管理体制对创新高效性和有效性的阻碍,转换管理模式实行流程再造,将数据的管理要素治理与工程管理模式相融合,相互支撑、形成互动,实现双向的顶层设计,建立截然不同的组织模式,实现有机融合,进而转变企业传统的管理模式。[10]

(二)分层次、逐步建立全面系统的大数据体系

构建政府、企业、个人的大数据系统,应从营造大数据发展环境的角度予以全面、系统化考虑,建立层次分明的大数据治理体系,包含数据的所有权、体制机制、共享、安全等多方面的内容,并且需要从制度法规、标准规范、应用、技术等多方面提供支撑。

未来几年数据泄露事件的增长率也许会达到 100%,除非数据在其源头就能够得到安全保障。数据安全是大数据应用必须要下大力气做好的一项必不可少的工作[11],从方案、措施到管理流程等每一个环节都要有规可依。从法律法规层面,应明确数据的所有权,将其作为数据流通、交易和保护的基础;制定促进数据共享开放的政策法规和标准规范,使企业作为大数据的拥有方与政府监管部门很好地配合,促进政府数据和企业数据的融合应用;并且出台数据安全与隐私保护的法律法规,保障国家、企业和个人的数据安全。

对于各不同行业,应对本行业发展的共同利益和战略规划进行充分考虑,建立相应的能够规范本行业的数据管理机构和管控制度,制定行业内的数据共享与开放的规则和技术规范[12],促进行业内数据的共享交换和融合应用。

对于企业而言,在数据治理体系建设中,大数据资源建设最重要的前提是数据的共享开放,重点是要提升企业对数据全生命周期的管理能力[13],在保障企业自身数据安全的前提下,促进各部门之间、不同企业之间的数据流通。同时,所有企业,无论规模大小,都需要从新的角度来确保企业自身以及客户的数据安全,所有数据在创建之初就需要获得安全保障,而并非在数据保存的最后一个环节。

(三)确保数据质量,防范大数据发展带来的风险

一方面,应当认真思考数据质量的管控与自动化数据动态采集、物联网、AI 等先进技术的充分结合,确保数据的准确性与真实性。另一方面,随着大数据的不断发展,可能会出现一系列新的风险,例如企业借助其行业优势,不断获取行业数据,极易出现数据垄断的现象,这种数据垄断不利于行业、企业的健康发展,甚至有可能威胁到国家的安全。建议国家出台相应法规,对大数据的应用予以规制,以防范相关风险的出现。

五、结语

随着油气生产步入“大数据”时代,明确大数据在石油工程应用中面临的挑战与发展思路至关重要。开展大数据处理技术的研究,突破技术瓶颈和障碍,更好地进行大数据的管理与挖掘,形成统一的、及时更新的、共享的地质与工程基础数据库,并进行科学分析整理与开发,对于石油工程大数据更好地应用于油田生产实践、提高油田生产决策的科学性具有重要的现实意义,同时也是石油企业提高核心竞争力的重要手段。

随着企业对“大数据”的认识和应用逐步深入,对数据的理解和把握也更加具体。大数据技术为油气行业的发展提供了新的机遇,提升对大数据的正确认知和具备“大数据思维”是促进数据有效共享和高效融合的前提[14],也是油气行业大数据平台建设的重要保障。此外,油气行业大数据在发展先进信息技术的同时,应广泛促进与物联网、AI 等技术的融合;同时加快培养适应未来需求的合格人才,助力提升油气勘探开发的能力和水平,提升数字时代我国油气行业的综合竞争力。

参考文献:

[1] 陈中普,王长在,任立春,等.油气田钻完井大数据技术研究与应用前景展望[J].录井工程,2018,29(4):1-6.

[2] 李金诺.浅谈石油行业大数据的发展趋势[J].价值工程,2013,29(11):172-174.

[3] 檀朝东,陈见成,刘志海,等.大数据挖掘技术在石油工程中的应用前景展望[J].石油工程技术,2015(1):49-51.
[4] 李国欣,王峰,皮学军,等.非常规油气藏地质工程一体化数据优化应用的思考与建议[J].中国石油勘探,2019,24(2):147-152.
[5] 林东龙,周大通,戴倚霞,等.利用大数据技术开发全球能源信息系统——中石油的实践[J].价值工程,2019,15(2):24-29.
[6] 郑军,尹兆涛.中国石油应对"大数据"的策略分析[J].石油规划设计,2013,24(6):27-29.
[7] 徐远震.科技兴油 挺起国家能源安全的脊梁[N].中国石油报,2020-02-21(04).
[8] 郑兴扬,张彤,刘昊伟.大数据分析在国际石油贸易领域的应用现状与发展趋势[J].石油科技论坛,2018(3):41-46.
[9] 深刻解析工业大数据:创新应用的挑战不断、40 年工业大数据研究总结工业智造关键 5 要素[EB/OL].(2019-02-12)[2020-02-25]. http://www.ciotimes.com/manufacturing/168932.html/.
[10] 刘合.采油工程持续融合创新管理与实践[J].石油科技论坛,2018(5):1-7.
[11] 何沙,胡尧,王林元,等.基于大数据技术的应急决策指挥体系构建——以中国国际石油合作突发事件为例[J].科技管理研究,2017,28(3):163-168.
[12] 梅宏.大数据发展现状与未来趋势[J].交通运输研究,2019(5):1-11.
[13] 刘合,孟思炜,苏健,等.对中国页岩气压裂工程技术发展和工程管理的思考与建议[J].天然气工业,2019,39(4):1-7.
[14] 刘合,金旭,丁彬.纳米技术在石油勘探开发领域的应用[J].石油勘探与开发,2016,43(6):1014-1021.

责任编辑:曲 红

Challenge and Development of Big Data Application in Petroleum Engineering

SU Jian[1,2], LIU He[1]

(1. *PetroChina Research Institute of Petroleum Exploration & Development*, *Beijing* 100083, *China*;
2. *School of Economics and Management*, *China University of Petroleum*(*East China*), *Qingdao*, *Shandong* 266580, *China*)

Abstract: In recent years, with the deepening of domestic oil and gas exploration and development, a large number of oilfield data resources have been accumulated in geology, logging, geophysical prospecting, exploitation and other links, and the informatization, digitalization and intelligent constructions are in progress. Oil and gas exploration and development has entered the era of "big data". As a link of multi-disciplinary integration, big data is of great significance to petroleum engineering. However, due to its huge data volume, complex structure, diverse forms and various types, there is no unified, timely updated and shared basic database of geology and engineering of oil and gas wells, and there is a lack of classification, sorting and secondary development of data. This paper sorts out the challenges faced by the application and integration of big data in petroleum engineering, and puts forward the data processing methods and development ideas, which has certain guiding significance for the better application of big data in oil field production practice and improving the scientific decision-making of oil field production.

Key words: big data; petroleum engineering; key technologies; management mode