

普光智能气田岗位数据流分析方法

陈惟国¹ 易建锋¹ 张捍卫² 付培晨¹ 李涛² 杨晓燕²

1. 中国石化中原油田分公司 2. 中国石化中原油田普光分公司

摘 要 岗位数据采集是数据资源中心建设的重要组成部分之一且直接关系到后者的质量和水平, 普光气田数据采集过程中存在着各专业数据库异构建设、数据无法共享、标准不一、多头采集、冗余存储、管理困难等问题。为了实现面向生产岗位、全业务链提供数据服务和应用, 在岗位中实现数据采、存、管、用一体化的目标, 基于中国石化勘探开发数据模型及源头数据结构设计, 结合普光气田的实际业务, 分析岗位业务关系、数据流业务流程和业务数据元素, 以紧密相关的一组数据作为数据集合, 分析数据的传递和相互间的作用, 建立了满足普光智能气田建设要求的岗位数据流分析方法, 形成了一套较为完善的资产化数据资源管理体系, 并进行了现场应用验证。结论认为: ①较之于传统的数据分析方法, 普光智能气田岗位业务数据流分析方法创新了数据流向分析, 以业务流程为驱动、岗位场景为验证, 实现了岗位数据的统一采集、集中存储、共享使用和统一管理, 取得了良好的应用效果; ②该研究成果提供了技术路线、可行方法和标准模板, 可推广复用。

关键词 普光气田 智能化 岗位 数据流 数据集 分析方法 编码 应用 采存管用一体化

DOI: 10.3787/j.issn.1000-0976.2018.10.006

A post-data flow analysis method for the Puguang Intelligent Gas-field Project

Chen Weiguo¹, Yi Jianfeng¹, Zhang Hanwei², Fu Peichen¹, Li Tao² & Yang Xiaoyan²

(1. Sinopec Zhongyuan Oilfield Company, Puyang, Henan 457000, China; 2. Puguang Branch of Sinopec Zhongyuan Oilfield Company, Dazhou, Sichuan 635000, China)

NATUR. GAS IND. VOLUME 38, ISSUE 10, pp.47-51, 10/25/2018. (ISSN 1000-0976; In Chinese)

Abstract: Post data collection is an important part in the construction of data resource center, and it is directly related to the construction quality and level. The data collection in the Puguang Gas Field is faced with heterogeneous construction, data non-sharing, different standard, multi-source collection, redundant storage and difficult management of different professional data bases. In order to provide data services and applications for production posts and full service chains and realize the integrated data collection, storage, management and utilization in the post, the post business relationship, data flow business process and business data elements were analyzed based on Sinopec's exploration and development data model and the structure design of source data, combined with actual business of the Puguang Gas Field. Then, the transmission and interaction of data were investigated with a closely related set of data as the data set. In this way, a post data flow analysis method which meets the construction requirements of the Puguang Intelligent Gas-field Project was established, and a complete set of asset-based data resource management system was developed. Finally, they were applied on site for verification. It is indicated that compared with the traditional data analysis methods, the post-data flow analysis method is more innovative with better application effect. It realizes unified collection, centralized storage, shared use and unified management of the post data with the business process as the driving force and the post scenario as the verification. These research results provide the technical route, feasible methods and standard templates and they are worth popularizing and applying.

Keywords: Puguang Gas Field; Intelligent; Post; Data flow; Data set; Analysis method; Coding; Application; Integration of data collection, storage, management and utilization

基金项目: 国家科技重大专项“高含硫气藏安全高效开发技术(三期)”(编号: 2016ZX05017)。

作者简介: 陈惟国, 1963年生, 教授级高级工程师, 现任中国石化中原油田分公司副总经理; 主要从事采油气工程等方面的研究和管理。地址: (457001) 河南省濮阳市中原路277号总经理办公室。ORCID: 0000-0003-3136-5216。E-mail: chenwg.zyyt@sinopec.com

0 引言

数据是企业的重要资产，已越来越成为社会的共识。数据是大数据分析、机器学习训练、智能分析预测的基础；数据既是价值的源泉，也是创新的基础。数据资源共享中心是普光智能气田建设的基础。普光气田业务岗位涵盖勘探、开发、集输、净化、安全环保、生产经营和管理等各个方面，业务流程长，涉及专业多，数据来源多，数据量大^[1-3]；主要数据来源包括静态（文档类）数据、SCADA/DCS实时数据、岗位采集数据。岗位数据采集是数据资源中心的重要组成部分。

在进行岗位数据采集和应用系统建设的过程中，首先进行了业务的梳理和调研，发现影响普光岗位业务应用的关键因素是数据流分析部分^[4]，主要存在5点不足：①数据无法全覆盖普光气田生产管理业务；②数据本身没有进行数据流分析，数据的流向不清楚，无法明确数据的采和用；③数据流程不明确，依赖各系统内部管理流程进行质量控制，造成多头采集，数据一致性差；④数据标准（各岗位）不统一，没有明确C/R关系，无法合理高效存储，实现数据唯一；⑤数据共享困难，跨系统数据难以交互，缺少统一的查询、管控、发布机制。

在普光智能气田建设规划顶层设计阶段，首先遇到的难题就是数据的规范化和标准化问题。普光智能气田的首要任务就是面向岗位的全业务链提供数据服务和应用，厘清数据的来源与去向，组成完整的数据链，建成数据资源共享中心。采取全新的面向岗位数据流分析方法，实现数据有源、流向有序、应用明确、环环相扣数据标准和流程，在岗位采集和应用中实现数据采集、存储、审核、应用一体化、规范化和标准化，为普光智能气田建设奠定坚实的数据资源基础。

1 岗位数据流分析方法研究

1.1 分析方法研究

通过对普光业务岗位预分析，自下而上地调研与找出存在问题，并结合普光气田实际进行分析研究，形成了一套适合油气田岗位数据流分析方法^[5]，编制了《岗位业务数据流分析指南》。

岗位数据流分析介于业务流程分析和数据元素分析之间^[6-7]，它以紧密相关的一组数据作为数据集，来分析数据的传递和相互间的作用。既要描述

原始数据的语义特征，又要描述数据集的整体特征。数据集分解的粒度大小^[8]，不仅关系到岗位数据整体准确表现，而且还为数据结构的提取和抽象提供便捷实现途径。

在普光工作流程、业务岗位分析的基础上，对岗位产生的资料进行结构化、标准化、规范化加工处理和优化整合，形成可用于建库的“普光气田数据模型标准”，成为数据模型和应用系统的关键桥梁。为普光一体化服务平台的数据采、存、管、用提供数据结构支持，满足普光业务应用的需要。

1.1.1 数据流分析模板设计

根据普光分公司岗位说明书中岗位职责、工作内容及工作标准、考核标准，预分析每个岗位使用和产生资料的信息^[6]，根据岗位定制调研模板并填写岗位描述工作模板部分内容，由下往上逐个岗位地梳理岗位之间的数据流向关系，设计岗位业务划分模板及数据集定义模板，提取可物理实现的数据字典，并记录资料样例，根据模型需求，为量纲设计阈值标识符号模板。利用模板对普光相关岗位涉及的业务流、数据流、数据关系、数据算法、数据量纲、数据阈值、数据定义进行数据和数据集分析^[7]，需求覆盖普光气田全业务域，针对性补充、完善、调整模板内容，为面向普光分公司岗位采集和应用，提供了翔实的信息和依据。

1.1.2 岗位业务数据集提取

普光气田一个岗位通常就是一个应用场景，即一个业务单元，对于操作岗可能有多个业务单元，每个业务单元产生/使用多张报表，每张报表可划分出一个或多个数据集。从岗位描述入手，参照岗位使用和产生资料，定义每个数据集来源和流向、频次，以数据集为单位有序流动并提取可物理实现的数据字典^[8]。数据集取之于报表，来源于岗位，满足岗位采集与应用的需要，数据集比较稳定，不因岗位变动而变化^[9]；满足边用边采，一次采集，共享使用。与普光智能气田项目业务划分成果进行对接，形成普光业务模型^[2]（图1）。

数据集包含岗位属性和时空特性，具备流向和类型特征，更加符合岗位特性。数据集划分应以不割裂岗位完整性为原则^[10]，按照不同时间维度、不同业务类型、不同设备装置等进行划分，同一岗位业务单元中数据集划分的维度保持一致^[8]；数据集粒度以数据的关联程度、岗位间便于传递为尺度，通过标准化的流程和工具对数据进行处理可以保证一

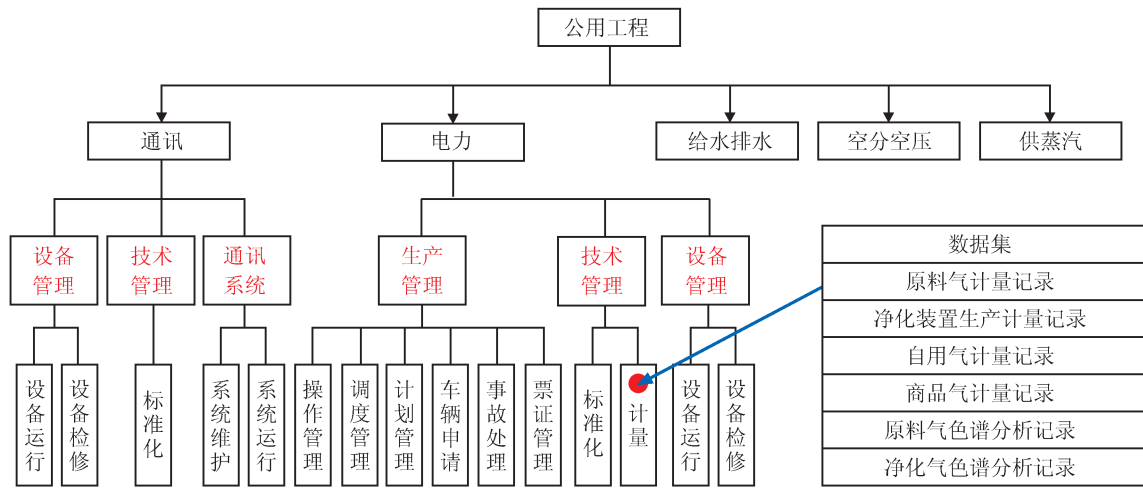


图 1 数据集与净化业务对接图

个预先定义好的高质量分析结果。

借鉴中石化勘探开发数据模型，结合普光气田业务实际，定义了 28 项数据集内容。数据集提取原则使用二维数据表结构，但为了满足普光业务数据三维或多维数据存储的规范化、标准化，需要增加附录代码表进行约束，对量纲标准化^[11]。数据集中数据项高聚集，数据关联性较强，数据单一；数据集组合可以实现普光各种岗位业务需求，与其他数据库融合较为容易；并可通过岗位交叉验证，保证数据的完整性和准确性；增加计算数据项描述，为应用提供指导。

1.1.3 数据模型设计

普光气田岗位分为基层操作岗、技术管理岗、综合管理岗等，一般来说基层操作岗面向的数据（报表）比较单一，它产生的数据来自各种监控、监视、

检测设备，向技术管理岗提供数据；技术管理岗一般制定、审核、平衡、汇总各类数据，向下（操作岗）发送各项指标、向上（综合管理岗）分发统计数据；综合管理岗根据统计数据，结合实物价值量，提供各种可供决策量化经济指标数据，供更高层次管理者决策。根据普光岗位业务划分模板和数据集定义模板，分析数据集的流向，利用数据项的 C/R 分析及验证，追溯数据在每个岗位中所处的生命周期状态（Creat、Read、Update、Delete）^[12]，从而实现岗位数据的采集、应用和监控。

通过对普光岗位业务流及数据流分析，编制数据集数据字典，建设数据模型^[12]，厘清了普光数据流转流程及岗位之间依赖关系^[13]，保证数据质量，为普光气田数据管理、数据流控制、数据应用打下了基础（图 2）。

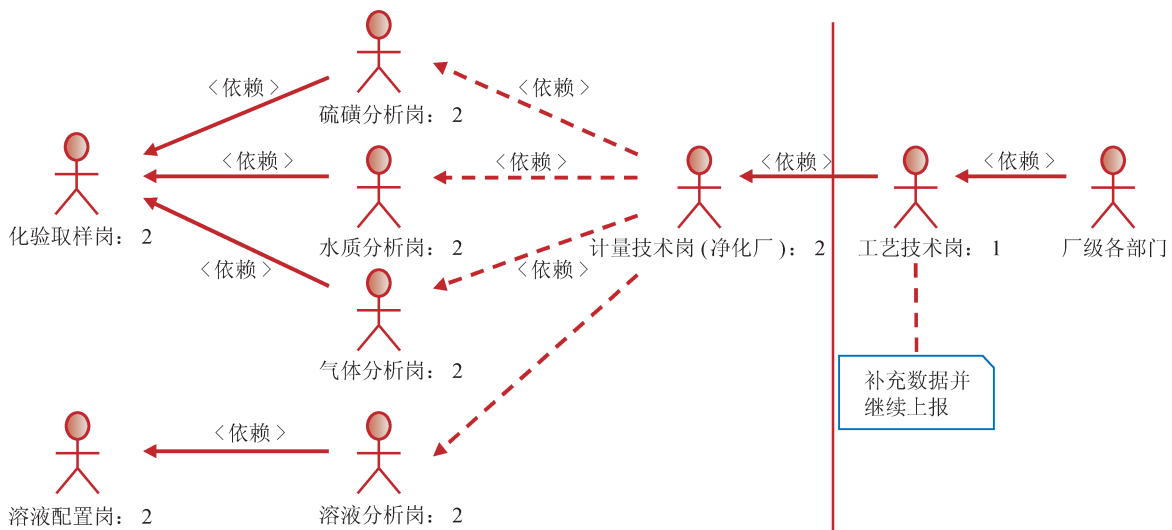


图 2 计量化验站部分岗位依赖图

通过对普光净化厂业务关系确定及数据报表的上下验证分析,可确定数据来源和流向,理顺了净化厂岗位之间数据关联及引用关系^[14],为数据唯一性、完整性、准确性打下了基础。

净化业务岗位数据流分析数据集,可以直接作为数据库结构来实现数据存储,也为数据存储优化提供了基础(图3)。

1.1.4 数据流分析编码规则制定

参照中石化勘探开发数据模型编码标准,制订了一套中国石化中原油田普光分公司(以下简称普光分公司)岗位名称、数据集名称、数据项名称编码规则。

岗位名称、岗位分类使用普光分公司岗位说明书的标准名称,岗位分类沿用普光分公司人力资源部设定的岗位分类。岗位编码和数据集编码兼顾系统性、通用性、可扩展性、可识别性。

以普光气田业务模型编码根据部门(二级)和单位(三级)设计,岗位编码采取11位编码,数据集采用15位编码(11位岗位码+4位数据集码)^[14]。

数据集作为岗位中最基本业务单元,不仅为独立完整数据表存储数据,也为数据应用、数据管理、数据传递搭建最基本的场景。数据集提取和命名原则上采用三段式方法:数据核心对象(可含数据所属组织机构)+时间维度(频度)+数据类型(动宾结构)^[9]。

数据项是数据集中不可再分基本元素。数据项代码可由数据项中文名汉语拼音首字母组成,字母相同则尾部可后缀一个数字以示区别。数据项代码避免使用SQL语句默认字符串;数据项代码最长位数在10位以内。

1.2 分析方法应用

岗位业务数据流分析方法在普光智能气田各岗位数据流分析中全面应用,基于对普光采气厂、净化厂、应急救援中心、生产服务中心及外包生产单位254个关键岗位的业务和数据流程进行梳理和分析。通过业务数据建模与标准化,共提取到数据集611个,提取附录代码表74张,标准化量纲134个。并在普光气田岗位数据采集及应用系统中部署实施,得到验证,见到了较好的效果。

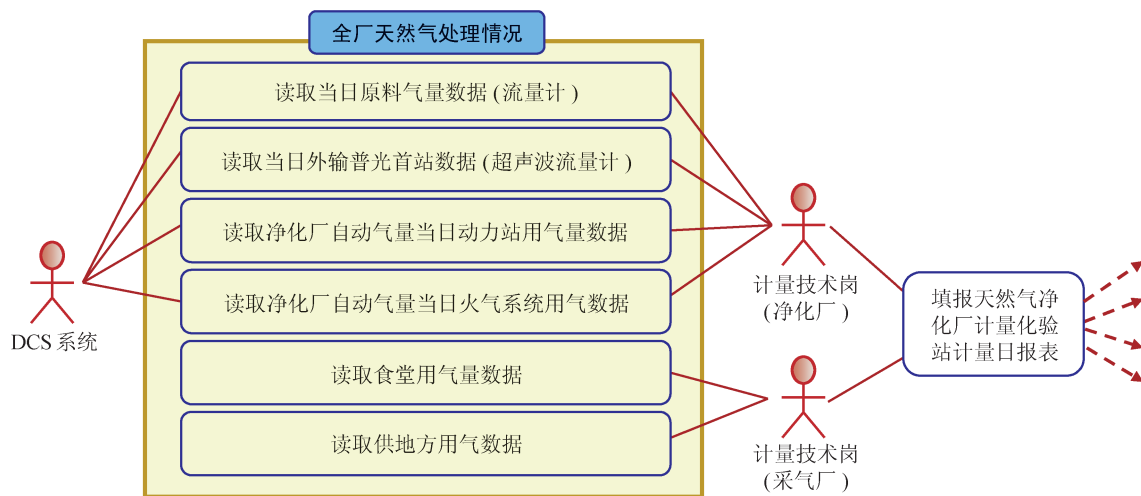


图3 计量化验站岗位数据流图^[12-13]

2 现场应用效果分析

通过普光气田岗位数据采集及应用系统,实现基层岗位的数据采集、岗位数据计算处理、汇总、岗位报表生成、岗位数据查询等功能。每个岗位具有独特应用场景,通过业务流程分析和定制,按照原有工作模式,提交表格完成数据采集和入库。由原来的逐级录入转变为固定模板计算机处理高效流程,建立了数据一次采集支持多级应用的功能。改变了原来多种模板录入、定时抄取数据、邮件上报后的重

复粘贴、人工计算汇总的工作方式,解决了数据共享的及时性和准确性问题,同时减轻了基层岗位人员数据上报的负担,杜绝了数据重复录入,提高了基层生产运行和协同效率。岗位数据采集基本实现普光气田岗位基础数据采集数字化、报表生成自动化。

以普光天然气净化厂计量日报为例,计量化验站的化验取样岗首先提交取样确认单报表至气体分析技术岗,录入天然气组成测定原始记录报表;计量交接岗录入净化厂供采气厂用气量报表,净化车间的工艺技术岗录入净化车间天然气处理情况报表,

数据都提交至计量技术岗进行审核汇总,最后提交至技术办的工艺技术岗进行审核并使用。

通过对普光气田岗位数据流统一分析,建立了普光气田数据流分析方法和数据模型,覆盖了普光气田勘探、开发、净化、生产保障全业务链的岗位,为数据的标准化和数据传输转换提供了依据。岗位应用系统的数据采集录入直接进入普光中心数据库,作为业务数据进行存储。数据流分析为数据的完整性、齐全性、唯一性、安全预警应用打下了坚实基础。

3 结束语

普光智能气田岗位业务数据流分析方法研究成果,为实现普光气田各部门、直属单位统一的数据分析、数据建模、数据采集、数据应用提供了技术路线、可行方法和标准模板。满足面向岗位数据采集、应用、管理的需求,实现数据的完整闭合,交叉验证,为今后油气田的岗位数据分析工作积累了宝贵的经验,有力推动了信息化与业务的深度融合。

普光智能气田业务岗位数据流分析方法,补充了混合天然气勘探开发业务模型,完善了中石化业务模型。可在类似油气田进行推广复用,并为智能气田建设的主数据管理和岗位应用建设提供借鉴。

参 考 文 献

- [1] 张世民,彭鑫岭.普光气田开发(优化)方案[R].达州:中国石化中原油田普光公司,2008.
Zhang Shimin & Peng Xinling. Development (optimization) program for Puguang Gas Field[R]. Dazhou: Sinopec Zhongyuan Oilfield Puguang Company, 2008.
- [2] 何生厚,曹耀峰.普光高酸性气田开发[M].北京:中国石化出版社,2010.
He Shenghou & Cao Yaofeng. Development of Puguang High Sour Gas Field[M]. Beijing: China Petrochemical Press, 2010.
- [3] 彭鑫岭,张世民,张雪松,李继强.普光气田气井投产层段优化方法及效果[J].天然气工业,2011,31(5):61-63.
Peng Xinling, Zhang Shimin, Zhang Xuesong & Li Jiqiang. An optimization method for the production pay zone of a gas well and its performance in the Puguang Gas Field[J]. Natural Gas Industry, 2011, 31(5): 61-63.
- [4] 李岩,王惠文,叶明.数据流分析与技术研究[J].计算机工程与应用,2008,44(15):8-11.
Li Yan, Wang Huiwen & Ye Ming. Data stream analysis and research[J]. Computer Engineering and Applications, 2008, 44(15): 8-11.
- [5] 汪小飞,赵克佳,田祖伟.数据流分析的关键技术研究[J].计

算机科学,2005,32(12):91-93.

Wang Xiaofei, Zhao Kejia & Tian Zuwei. Research of key techniques in data flow analysis[J]. Computer Science, 2005, 32(12): 91-93.

- [6] 马全林,何庭全,郭晓军.利用要素记点法进行岗位价值评估[J].科技与企业,2015(23):7-48.
Ma Quanlin, He Tingquan & Guo Xiaojun. Evaluation of post value by means of element recording[J]. Science-Technology Enterprise, 2015(23): 7-48.
- [7] 李念念,许清平,王卫栋.基于IEEE1232的故障树诊断模型研究[J].计算机测量与控制,2014,22(11):3457-3459.
Li Niannian, Xu Qingping & Wang Weidong. Research on fault tree diagnosis model based on IEEE1232[J]. Computer Measurement & Control, 2014, 22(11): 3457-3459.
- [8] 于万钧,冯鹏飞.抽象网格工作流中数据集划分方法的研究[J].化工自动化及仪表,2013,40(8):1012-1015.
Yu Wanjun & Feng Pengfei. Research of datasets division in abstract grid workflow[J]. Control and Instruments in Chemical Industry, 2013, 40(8): 1012-1015.
- [9] 梁竹.支持数据挖掘算法选择的数据集特征提取研究[J].数字技术与应用,2011(4):117-120.
Liang Zhu. Dataset character extracting research to support[J]. Digital Technology and Application, 2011(4): 117-120.
- [10] 孙雁,田忠和,王泉德.一种高效的数据集划分关联规则挖掘算法[J].计算机工程,2002,28(12):118-120.
Sun Yan, Tian Zhonghe & Wang Quande. An algorithm of dataset-partition association rules mining[J]. Computer Engineering, 2002, 28(12): 118-120.
- [11] 周理,陈赓良,郭开华.对天然气推荐性国家标准GB/T 31253的讨论[J].天然气工业,2017,37(12):87-91.
Zhou Li, Chen Gengliang & Guo Kaihua. A discussion on the Chinese national standard GB/T 31253: From Recommendation (GB/T) to Guidance (GB/Z)[J]. Natural Gas Industry, 2017, 37(12): 87-91.
- [12] 黎升洪,蔡飞.面向对象数据库设计方法[J].计算机与现代化,1999(1):13-15.
Li Shenghong & Cai Fei. Subject-oriented database design[J]. Computer and Modernization, 1999(1): 13-15.
- [13] 徐天伟,夏幼明,李凯佳.数据流图描述语言的研究与设计[J].云南师范大学学报(自然科学版),2003,23(3):4-8.
Xu Tianwei, Xia Youming & Li Kaijia. Research and design on data flow graph description language[J]. Journal of Yunnan Normal University (Natural Sciences Edition), 2003, 23(3):4-8.
- [14] 田瑞兰.试谈主数据管理信息系统的建设[J].电脑编程技巧与维护,2011(18):77-79.
Tian Ruilan. On the master data management information system construction[J]. Computer Programming Skills & Maintenance, 2011(18): 77-79.

(修改回稿日期 2018-07-25 编辑 韩晓渝)