文章编号:1674-8190(2024)03-071-10

## 基于大数据分析方法的商用飞机 MOC9 试验项目 管理可视化模型研究

于海燕1,余汇2

(1.上海飞机设计研究院 项目中心,上海 201210)

(2.上海飞机设计研究院飞机电控集成工程技术所,上海 201210)

摘 要:实践表明,大数据技术可以对管理过程赋能,商用飞机MOC9试验项目作为商用飞机研制复杂高端项目群中的重要项目,其在研制期间会产生大量项目管理数据,因此,研究如何利用大数据技术推动该类项目实现项目目标十分必要。基于商用飞机研制项目群下MOC9试验项目的特点及其对大数据管理的需求,结合大数据技术处理流程和方法对MOC9试验项目从架构策划、数据获取与存储、数据清洗处理、分析挖掘、可视化展示多方面开展研究。结果表明:通过大数据技术对MOC9试验项目建立量化的可视化模型,能够对项目经理掌握项目整体情况、定位问题根本原因、识别项目潜在风险提供必要的帮助。

关键词: 商用飞机; MOC9试验项目; 大数据; 可视化; 项目监控

中图分类号: V216; V217; TP311.13; C931.9

文献标识码:A

**DOI:** 10. 16615/j. cnki. 1674-8190. 2024. 03. 07

# Study on visualization model of commercial aircraft MOC9 test project management based on big data analysis method

YU Haiyan<sup>1</sup>, YU Hui<sup>2</sup>

- (1. Program Center, Shanghai Aircraft Design and Research Institute, Shanghai 201210, China)
- (2. Aircraft Electronic Control Integration Engineering Technology Institute, Shanghai Aircraft
  Design and Research Institute, Shanghai 201210, China)

**Abstract:** Practice has shown that big data technology can empower the management process. As an important project in the complex high-end program of commercial aircraft development, the commercial aircraft MOC9 test project will generate a large amount of project management data during development. Therefore, it is necessary to study how to use big data technology to promote such projects to achieve project objectives. Based on the characteristics of the MOC9 test project under the commercial aircraft development program and its demand for big data management, this article conducts study on the MOC9 test project from multiple aspects such as architecture planning, data acquisition and storage, data cleaning and processing, analysis mining, and visual display, combined with big data technology processing processes and methods. The results show that establishing a quantitative visual model for MOC9 projects through big data technology can provide the necessary help for the project managers to master the overall situation of the project, locate the root causes of problems, and identify potential risks of the project

Key words: commercial aircraft; MOC9 test project; big data; visualization; project monitoring

收稿日期: 2023-03-30; 修回日期: 2023-05-21

通信作者:于海燕(1981-),女,学士,高级工程师。E-mail:yuhaiyan@comac.cc

引用格式:于海燕, 余汇. 基于大数据分析方法的商用飞机 MOC9 试验项目管理可视化模型研究[J]. 航空工程进展, 2024, 15(3): 71-80. YU Haiyan, YU Hui. Study on visualization model of commercial aircraft MOC9 test project management based on big data analysis method[J]. Advances in Aeronautical Science and Engineering, 2024, 15(3): 71-80. (in Chinese)

#### 0 引言

商用飞机研制项目是复杂高端项目群,飞机产品本身即使不把紧固件包含在零件数目以内,用各种材料制造的零件有几十万个<sup>[1]</sup>,项目生命周期长,从立项、设计、制造、试验、试飞到交付运营有时长达十几年,为了按时、按质量、按预算将飞机交付给用户,需要成千上万的人在复杂的组织结构下密切配合协同工作<sup>[1]</sup>,很多项目采用主制造商一供应商模式,项目涉及全世界上百个供应商,管理任务具有挑战性。

按照航空器型号合格审定程序要求,MOC9 试验是表明设备合格性、对设备进行鉴定的符合 性方法之一<sup>[2]</sup>。商用飞机机载设备通过 MOC9 试 验是取得型号合格证的必要条件。国内,谢春稳 等[3]论述了燃油系统产品的设备合格鉴定试验更 改流程和更改控制;章光灿[4]阐述了设备合格鉴定 试验各阶段的划分、具体内容和进出准则。大数 据技术研究方面,Li Sujie等[5]研究了通过建立大 数据平台每天收集多种数据信息以便飞机制造商 和航空公司优化民用飞机飞行; Vikas 等[6]介绍了 一种应用 Hadoop 的大数据项目监控软件,可以处 理项目的海量数据并可视化项目监控数据,对处 理多个相关和不相关的数据库进行分析; Kusek 等[7] 阐述了项目基础的大数据应用,研究了使用 Hadoop 进行数据可视化建立数据库的应用;魏壮 字等[8]提出对大数据量的民用飞机质量数据采用 基于星型有向图的交互式可视化分析方法,将关 联规则挖掘的整个过程通过大数据可视化的方式 展现出来;陈金等[9]对民用飞机运行大数据分析平 台整体系统架构进行研究;袁博等[10]从大数据分 析应用视角分析文本挖掘技术实现飞机装配质量 问题智能诊断的过程模型;张磊磊等[11]研究了建 立统一的数据资源管理平台采用大数据及可视化 分析工具实现对各测试环节测试数据的综合管 控。上述文献未关注商用飞机 MOC9 试验项目管 理可视化研究。

现有文献虽然对航空业质量类大数据应用、适合航空公司应用的大数据研究以及使用 Hadoop

对不同数据库进行分析有一定的研究,但未关注 使用大数据技术对商用飞机行业项目管理数据进 行研究。

商用飞机项目群管理难度极大,会产生海量项目管理数据,如果能合理应用大数据技术提高商用飞机项目群的项目管理效率,将产生巨大经济效益。因此,本文以商用飞机项目群下的MOC9试验项目为例,利用大数据技术对项目管理数据从开展架构策划、数据获取与存储、数据清洗处理、分析挖掘、可视化展示方面进行研究,提出对项目监控进行可视化的多个建议。

## 1 MOC9试验特点及对大数据管理 的需求

MOC9试验项目是商用飞机研制项目群中的一个项目,商用飞机机载设备通过 MOC9试验是保证机载设备满足飞行安全要求和使用环境要求的证明。一架商用飞机上的机载设备多达几百项,涉及到的供应商几十家,单个设备需实施的MOC9试验可达几十甚至上百项,因此针对单架飞机而言,所有机载设备需开展的试验项目可达上万项。MOC9试验项目的核心工作包括合同签署、鉴定试验大纲(QTP)批准、试验件到位、试验室到位、目击人员到位等,试验通常要求在较短时间内完成,因此有其他类似项目的历史数据作为参考对商用飞机 MOC9试验项目按进度、质量和成本完成会起到事半功倍的效果。

MOC9试验通常由供应商实施,关于供应商设计能力和MOC9试验策划实施的能力考核非常重要,选择合适的供应商就成功了一半。以某作动器鉴定试验为例,分析监控数据看出从2020年4月供应商开始编制鉴定试验大纲并准备试验件制造工作,试验件生产周期5个月,预计2020年9月底完成生产,实际试验件制造多次延迟,到2021年5月底才完成。经过鱼刺图(如图1所示)分析试验件多次延迟是由于供应商在民用飞机设计方面对适航要求理解不够深入,导致设计产品多次无法通过MOC9试验要求,MOC9试验的延迟对整个项目群带来不良影响。

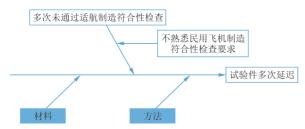


图 1 某作动器 MOC9 试验进度延迟鱼刺图 Fig. 1 Fishbone diagram of delay in the progress of an actuator MOC9 test

在新的商用飞机研制项目中,对该设备的供应商进行选择时需着重评估该供应商目前对适航要求的理解和落实能力。研究某卡箍产品,其实际进度比计划大幅延迟,通过查看 MOC9 试验的监控数据发现多次提到 QTP 修改、供应商不及时反馈,进一步挖掘原因后发现供应商配合度较低,合作意愿不强。原则上,在下一个项目初期选择供应商时需考虑该供应商在本次 MOC9 试验进程中的表现。

由于每个公司都有一定的人员流动率,完成一个项目后会有一定数量的人员流失,不能全靠人员存在于大脑或个人电脑的内容,因此这类项目管理监控数据非常重要,需要在项目策划阶段考虑项目管理大数据库的建设。

#### 2 大数据的概念和处理流程

大数据的发展已经经历了十余年,然而,研究者对大数据的认识不尽相同,而且也不断发生着变化<sup>[12]</sup>。麦肯锡对大数据的定义是:大数据是指无法在一定时间内用传统数据库软件工具对其内容进行采集、存储、管理和分析的数据集合<sup>[13]</sup>。数据类型分为结构化、半结构化、准结构化和非结构化<sup>[14]</sup>,如表1所示。

表 1 数据类型 Table 1 Data type

数据类型	数据内容	举例
结构化	包括预定义的数据类型、 格式和结构的数据	事务性数据和联机分析 处理
半结构化	具有可识别的模式并可以 解析的文本数据文件	自描述和具有定义模式 的 XML 数据文件
准结构化	具有不规则数据格式的文 本数据,通过使用工具可 以使之格式化	包含不一致的数据值和 格式的网站点击数据
非结构化	没有固定结构的数据,通 常将其保存成不同类型的 文档	TXT文本文档、PDF文档、图像和视频

大数据处理流程通常包括:数据获取与存储、数据处理、数据分析和数据可视化<sup>[15]</sup>,如图2所示。

- 1)数据获取与存储:目前行业有两种解释,一是数据从无到有的过程(WEB服务器打印的日志、自定义采集的日志等)叫做数据采集;二是把通过使用Flume等工具把数据采集到指定位置的这个过程叫做数据采集。
- 2) 数据处理:对采集到的数据进行格式整理、 数据清洗、过滤出脏数据等。
- 3)数据分析:根据需求开发分析语句,得出统计结果。
- 4)数据可视化:将分析所得的数据进行数据 可视化,一般通过图表进行展示。

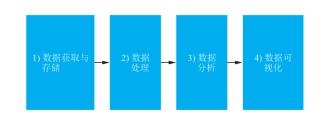


图 2 大数据处理流程 Fig. 2 Big data processing process

## 3 MOC9试验项目大数据技术应用

由于项目经理都会存储项目管理数据,因此对于在项目进行中才开始应用大数据技术的项目,项目经理可以直接把数据拷贝到数据仓库中。对于在项目策划阶段就打算应用大数据技术的项目,可以根据项目特点搭建数据平台,平台作为数据获取、存储和处理的载体。

#### 3.1 数据平台搭建需求

数据平台的搭建要考虑该项目的工作分解结构(WBS),如图 3 所示,可以看出: MOC9 试验项目的核心工作包括合同签署、鉴定试验大纲、试验件、试验室、目击人员到位等。根据工作流程可以把试验的主要阶段分为以下 10 个阶段(如图 4 所示)。

项目的利益攸关方包括主制造商的各专业部门、供应商、中国民用航空局(CAAC)等,如图 5 所示。

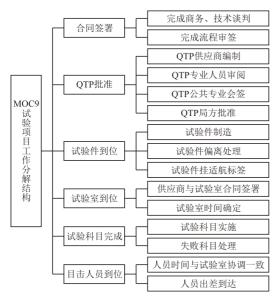


图 3 MOC9项目的工作分解结构 Fig. 3 MOC9 test WBS

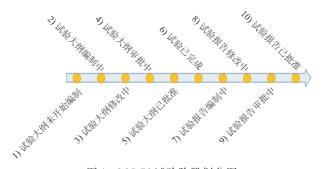


图 4 MOC9 试验阶段划分图 Fig. 4 MOC9 test stage division diagram



图 5 MOC9 试验项目的利益攸关方 Fig. 5 MOC9 test project stakeholders

数据平台设计既要考虑项目监控自身的需求 也要考虑不同利益攸关方,比如不同层级管理人员、CAAC代表、各部门对该项目中人员表现的需求,因此在平台设计时可以包含以上信息。通常只是记录试验状态无法满足对项目风险问题识别和应对的需求,因此可以增加"详细状态"列(如图6所示),记录各要素的状态、问题和行动项。

试验任务量↑	+ 负责 IPT 团队 <b>画</b>	+ 负责部门 <b>雪</b>	+ 试验责任人	+ 试验状态 <b>■</b> ↑	+ QTP 状态 <b>■</b>	详细状态
+ A						
+ B						
+ C						

图 6 MOC9试验项目数据平台设计 Fig. 6 MOC9 test project data platform design

#### 3.2 数据获取和存储

数据获取主要来源于项目的计划、设备供应商提供的进展以及各类会议,项目通常每周有例会进行监控状态、识别问题和风险,会议记录可以直接在平台上填报形成基础数据库。同时,可以使用Flume等工具收集与MOC9试验相关的其他子项目、技术专题和管理专题、各集成产品开发(IPT)团队例会等的纪要进行数据采集并处理分析后提取对MOC9试验项目有用的数据进行数据库完善。

另外,也可以收集供应商 MOC9 试验管理的相关历史数据,形成供应商 MOC9 试验项目管理能力的基础评价。该数据可以通过会议交流或者问卷调查的方法获取,问卷调查表如表2所示。

如果历史数据中计划完成时间和实际完成时间差距较大,说明供应商 MOC9 试验项目管理能力存在提升空间,可以在可视化阶段进行设置标出不同颜色以便重点关注。

表2 供应商MOC9试验管理历史数据问卷调查表

Table 2 Supplier MOC9 test management history data questionnaire

序号 供应商名称 是否有相关经验 设备名称 QTP批准计划 QTP批准实际时间 试验计划 试验实际完成时间 延迟原因

#### 3.3 数据处理

对于在平台上收集的数据,由于是按结构化进行的设计(参见表1所示的数据类型分类),基本不用进行数据清洗等工作可以直接使用。对于从

其他平台和其他子项目采集的数据,部分是会议 纪要等文本类非结构化数据,部分是公司ERP平 台上与QTP驳回、批准或试验件设计更改单、偏离 单处理等半结构化数据,需要进行数据清洗和 整理。

#### 3.4 数据挖掘分析

对于 MOC9 试验各 IPT 团队的总数量、各团队、各部门和各供应商的"试验状态"等数据都是在平台上按结构化设计填报的,因此不需要进行数据挖掘,直接提取即可满足项目监控基本需求。

每次例会会把监控到的详细状态、出现的问题和识别的风险填报到平台的"详细状态"栏目中以便采取措施(如图6所示),会议会形成行动项来推动项目按目标开展。上述问题类和风险类数据通常以文本形式描述,比如"流体敏感性试验故障,正在开展故障分析","详细状态"无法直接过滤统计出每一类问题和风险的数量,如果想获取每一类问题和风险的数量则需要管理人员逐项分析。为了提高效率,可以通过大数据挖掘中的"分类"方法对监控数据进行自动分类,如表3所示。

大数据挖掘中分类是常用方法,通过历史数据建立训练集,可以通过贝叶斯模型进行分类[16-17]。贝叶斯定理公式为: $P(A|B)=P(B|A)\cdot P(A)/P(B)$ 。

表 3 MOC9 试验监控数据分类 Table 3 MOC9 test monitoring data classification

序号	问题类别	风险类别
1	合同签署	合同签署
2	QTP	QTP
3	试验件	试验件
4	试验室	试验室
5	目击人员到位	目击人员到位
6	试验科目失败	试验科目失败

根据建立的模型,输入"详细状态"数据,模型可以批量分析设备目前是否有潜在风险,识别出风险类别或问题类别(如表3所示),下一个型号也可以使用历史模型进行风险预测和应对。

为了实现量化的问题识别和风险预测,可以对 WBS中的各项工作的阶段进行细分,在数据仓库设计时增加 QTP、试验件、试验室、试验等各要素,并把这些要素进行阶段细分,各细分阶段对应1个识别号,如表4所示。

表4 MOC9 试验项目数据仓库设计要素 Table 4 MOC9 Test project data warehouse design

QTP状态/问题	QTP识 别号	试验件状态	试验件 识别号	试验状态/问题	试验识 别号
供应商未开始编制 QTP,计划 X月 X 日开始编制	0	试验件未开始加工,计划于 X 月 X 日 完成加工	0	试验未开始,计划 X 月 X 日开 始试验	0
供应商在编制QTP,计划X月X日提交,仍未提交	1	试验件已开始加工,进行顺利,计划于 X月X日完成加工	1	试验已开始,进行顺利,计划 X月X日完成试验	1
供应商在X月X日已提交QTP	2	试验件已开始加工,出现问题	2	试验已开始,XX出现问题	2
主制造商负责专业在审阅,计划 X 月 X 日完成审阅	3	试验件出现偏离,供应商在编制偏离 报告	3	在进行XX问题根因分析	3
主制造商负责专业在 X 月 X 日完成 审阅	4	偏离报告已批准	4	XX科目已完成,计划开展YY 科目	4
主制造商公共专业在会签,计划 X 月 X 日完成会签	5	试验件于X月X日完成加工	5	试验已完成	5
主制造商公共专业在X日完成会签	6				
供应商在修改,计划 X 月 X 日提交	7				
QTP已在X月X日给局方代表汇报,局方代表提出修改意见	8				
QTP已在X月X日上审批流程	9				
QTP已被批准	10				

每次例会时不仅收集试验整体状态,还单独 收集QTP、试验件、试验室、出国团组批件到位、 目击人员到位、合同签署、试验科目和其他的细分 状态,根据细分状态的识别号可以按时间维度进 行可视化,制约项目进展的根本原因就变得显而 易见。

#### 3.5 数据可视化

可视化报表设计需要考虑各层级管理人员和各利益攸关方的管理需求,可以按IPT团队统计(如图7所示)、按部门统计、按项目阶段统计、按时间维度统计形成各种图表,也可以对IPT团队下属各工作包的状态进行"钻取"。

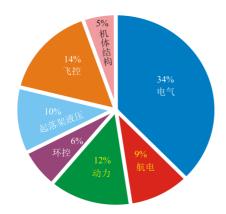


图 7 MOC9 试验可视化报表中 IPT 团队统计表 Fig. 7 Statistical table of IPT team in MOC9 test visual report

数据可视化还可以根据项目特色进行定制化设计,MOC9试验项目中QTP是所有工作的前置,但是难度也比较大,经常成为该项目的"关键路径"。有时QTP在供应商、主制造商内部多部门和

局方之间多次沟通修改,有时由于试验环境或方案更改导致QTP批准需要几个月,在管理层想梳理清楚问题的根源以便采取有力措施时,经常主制造商各专业和供应商都会提出对方反馈不及时影响进度,没有一个量化的工具可以一目了然地展现出各环节使用的时间和出现的问题。

如果在平台上按表4所示把各要素的细分阶段的状态数据进行结构化设计,每个状态都赋予特定的数值,这样可以以时间维度自动生成各要素状态图,在平台上可以看出QTP编制批准等工作的研制历程(如图8所示)。

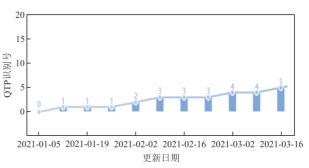


图 8 MOC9 试验 QTP 编制历程可视化图 Fig. 8 MOC9 test QTP preparation history visualization chart

如果数字1(表4中QTP识别号1的含义为"供应商在编制QTP,计划X月X日提交,仍未提交")持续很久,说明供应商未按计划提交QTP,延迟很久,就可以转到采购团队对应层级进行解决。还可以设置如果某个数字持续时间超过X周,系统提醒对应人员关注(如图9所示),避免等管理层关注时已经过去几个月,可快速形成可量化的风险预测结果。



图 9 系统提醒设置举例 Fig. 9 System alarm set-up example

可以在系统中开发风险自动识别功能,将历史数据调整后的值设置为基线,项目经理更新数据状态后,系统会自动匹配更新的数据与基线的关系,如果数据在基线范围内,显示绿色指示灯,如果已超过基线,显示黄色指示灯,如果超出一定范围,显示红色指示灯(如图 10 所示)。

序号	试验任务	状态监控灯	试验件 监控灯	试验室 监控灯	外协合同 监控灯
1	VSU		•		•
2	GIPC	•			•

图 10 指示灯设置举例

Fig. 10 Indicator light set-up example

考虑到供应商的综合能力对试验按要求完成的影响比较大,可以设置一套针对供应商 MOC9

试验管理能力的评价规则和基准,进行要素和权重的设置(如表5所示),对所有设备的供应商MOC9试验管理能力进行评估后与基准进行对比,把对比结果进行可视化展示(通常使用雷达图),当低于标准一定数值时进行风险预警提示。表5中"按时完成率"和"配合及响应程度"基准可以设置为90%~100%,"项目管理能力"和"适航能力"可以设置为80%~100%,4项因素权重相加为100%。

表 5 供应商 MOC9 试验管理能力评价要素权重表 Table 5 Supplier MOC9 test management ability assessment table

要素	实际	基准/%	得分	权重/%	加权得分
按时完成率		90~100		30	
配合及响应程度		90~100		25	
项目管理能力		80~100		20	
适航能力		80~100		25	

供应商 MOC9 试验管理能力低于标准较多,则需要在可视化展示平台上做出标注。

#### 4 应用实例

某商用飞机 MOC9 试验项目,具有试验科目 多、试验难度大、涉及单位和人员复杂等特点,其按项目进度、质量和成本完成对整个商用飞机项目群的要求取证有很大的影响。通过介绍实施大数据技术,项目经理可以高效掌握项目整体情况、定位问题根本原因、识别项目潜在风险。

## 4.1 某商用飞机 MOC9 试验项目数据平台 搭建需求

该项目的工作分解结构(WBS)如图 3 所示,工作流程如图 4 所示,项目的利益攸关方如图 5 所示,同时考虑到目击试验的实施需要 CAAC 代表、供应商、各专业人员的密切配合才能完成,CAAC 人力资源相对紧张,目击试验按时完成难度系数较高,因此增加了"目击试验"相关内容,以便对目击试验重点关注,对数据平台进行搭建,如图 11 所示,其中 MOC9 试验报告(QTR)、采购团队负责人、QTP文件号等多种细分栏目由于篇幅原因未完全展示。

序号	试验任	目击项 状态 ▼	负责IPT 团队 ▼			适航负 责人	试验状态		试验计划 完成时间 ▼	试验实 际完成 日期 🔻	详细试验状态 🔻
1	XX线缆 装置 MOC9 试验	未完成	电气系 统集成 IPT	UXX	魏X	姚X	QTP 已 批准	已发布	######		xxx
	XX板 MOC9 试验	已完成	电气系 统集成 IPT	广州 XX公 司	杨X	沈X	试验已 完成	已发布	######	#####	XXX

图 11 某商用飞机 MOC9 试验项目数据平台设计 Fig. 11 XX commercial aircraft MOC9 test project data platform design

## 4.2 某商用飞机 MOC9试验项目数据获取 和存储

该项目的数据获取主要来源于项目进度计划 和监控数据。进度计划部分包括年度计划和设备 试验详细计划(如图12所示)。

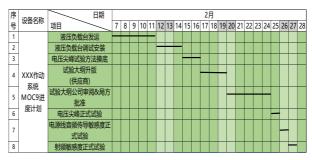


图 12 设备试验详细计划举例

Fig. 12 Equipment test detailed schedule example

监控数据包括设备供应商提供的进展(包括会议纪要、邮件、工程更改请求等)、项目例会纪要、工具收集QTP批准进展(收集工具如图13所示、收集结果如图14所示)、总师例会、项目总经理例会、局方适航审查会、1级IPT团队例会中与MOC9试验相关的决策。项目经理带领管理团队对以上数据进行采集后提取对MOC9试验项目有用的数据进行数据库完善。

名称	修改日期	类型	大小
chromedriver.exe	2018/3/2 0:02	应用程序	6,256 KB
getIDEAL.exe	2020/11/23 16:07	应用程序	8,207 KB
IDEAL技术文件查询模板.xlsx	2021/3/17 14:29	Microsoft Excel	17 KB
Readme.txt	2020/10/28 15:54	文本文档	1 KB
提取最后日期.xlsx	2020/11/24 14:05	Microsoft Excel	17 KB

图 13 数据收集工具 Fig. 13 Data collecting tool

席号	男要查的文件号	编号	名称	版本	状态	上次修改时间	创建时间
1			XX机气源系统预冷器	A. 3		2021/03/15 18:19	
2	E-C229SY1xx	E-C229SY1	XX飞机液压能源系统	A. 14	重新提交	2021/02/26 10:11	2020/02/21 0
3	E-C229SY1xx	E-C229SY1	XX飞机液压能源系统	A. 7	已发布	2021/02/26 10:09	2020/02/21 0
4	E-C224SY1xx	E-C224SY1	XX飞机外电源插座可	A. 1	审核	2021/02/22 09:35	2021/01/28 1
5	E-C224SY9xx	E-C224SY9	XX飞机主蓄电池充电	A.3	驳回	2021/02/20 16:59	2021/01/22 1

图 14 QTP状态收集结果 Fig. 14 Collected QTP status

#### 4.3 某商用飞机 MOC9试验项目数据处理

年度计划直接导入平台主界面,设备试验详 细计划导入到对应模块。收集的QTP批准状态数 据(如图 14 所示)更新到平台上"QTP批准状态" 栏目(如图11所示),试验件设计更改单、偏离单处 理等半结构化数据需要项目管理团队人员把相关 进展更新到"详细试验状态"栏目中,并标明日期。 各种会议纪要等文本类数据更新到"详细试验状 态"栏目中。

#### 4.4 某商用飞机 MOC9 试验项目数据挖掘

项目经理需要对整个项目的问题和风险进行 整体把控并采取措施。对问题和风险进行分类, 考虑到完成出国团组批件审批是出国目击的一项 重要前置,实际出现问题较多,涉及到公司总部等 多个相关部门,协调难度较大,因此作为一个单独 类型,具体分类方式如表6所示。

为提高效率,通过大数据挖掘中的"分类"方 法对"详细试验状态"栏目中的监控数据进行自动 分类。首先使用历史数据建立训练集,如表7所示 (由于篇幅原因截取部分训练集内容),其中"类 别"的数字与表6中序号和类别一一对应。

使用贝叶斯模型原理和 Python 软件建立分类 模型,通过训练后该分类模型可以自动识别平台 上"详细试验状态"栏目中以及各类会议纪要中的 文本类数据所属的问题和风险分类,某商用飞机 MOC9 试验文本数据分类程序如图 15 所示,程序 运行后,可以在表格中标识出该问题/风险所属类 别,如表8所示(由于篇幅有限只列出图15中程序 的前7项内容)。

表6 某商用飞机 MOC9 试验问题和风险分类 Table 6 XX commercial aircraft MOC9 test problem and risk classification

序号	问题类别	风险类别
1	QTP	QTP
2	试验件	试验件
3	试验室	试验室
4	出国团组	出国团组
5	目击人员到位	目击人员到位
6	合同签署	合同签署
7	试验科目失败	试验科目失败

表7 某商用飞机 MOC9 试验分类训练集 Table 7 XX commercial aircraft MOC9 test data classification training set

		classification training set
类别	风险/问题	字段
7	问题	试验多次失败,由于疫情物流停止试验设备 刚到位,预计2月疫情结束可以重新试验。
5	风险	由于疫情无法去现场目击,在协调远程目击
4	问题	未报出国计划,需要走计划外流程,周期较长,无法在3月出发
6	风险	项目已决策增加一个试验件,正在变更合同
6	风险	目前计划 2022年2月完成试验,如新增一个试验件可提前到 2022年1月,涉及商务待领导决策
2	问题	1号试验件下周完成耐久试验50%,试验件 需要返修
2	问题	试验件制造符合性检查未通过,下周四重 新做
1	风险	升版QTP可以下周上流程,预计1周可以 批准
3	风险	试验因试验室档期推迟到9月之后
1	风险	QTP说明材料供应商未反馈
	•••	

D:\anaconda\lib\site-packages\sklearn\base.py:299: UserWarning: Trying to unpickle e ing version 1.2.1. This might lead to breaking code or invalid results. Use at your https://scikit-learn.org/stable/model\_persistence.html#security-maintainability-limi warnings.warn(

D:\anaconda\lib\site-packages\sklearn\base.py:299: UserWarning: Trying to unpickle esusing version 1.2.1. This might lead to breaking code or invalid results. Use at your https://scikit-learn.org/stable/model\_persistence.html#security-maintainability-limi warnings.warn(

Building prefix dict from the default dictionary ...
Loading model from cache C:\Users\86150\AppData\Local\Temp\jieba.cache

Loading model cost 1.076 seconds.

Prefix dict has been built successfully.

[2. 7. 2.

图 15 某商用飞机 MOC9 试验文本数据分类程序

Fig. 15 XX commercial aircraft MOC9 test text data classification procedure

表 8 某商用飞机 MOC9 试验分类结果
Table 8 XX commercial aircraft MOC9 test data
classification result set

类别	风险/问题	字段
2	风险	试验件制造符合性检查的RFC填报供应 商和局方意见不一致,需要协调供应商
7	风险	供应商计划推迟至3月中旬完成试验
7	问题	新构型试验结果有部分问题,局方让重做
7	风险	后续要做砂尘和压力需公司批让步单
6	风险	在进行合同商务谈判,按流程合同签署前 供应商不会提前工作
6	问题	A4合同仍未确认文本,后续工作供应商 不开展
2	问题	试验件出现偏离,供应商在编制偏离单

## 4.5 某商用飞机 MOC9 试验项目数据可 视化

对于在平台上收集的数据,可以直接使用,按管理需求生成各类可视化图表(实际图表较多,篇幅有限不一一列举),比如QTP批准状态按IPT团队的分布图(如图16所示)、年度计划按月份的分布图、实际完成时间与年度计划对比的燃尽图等。

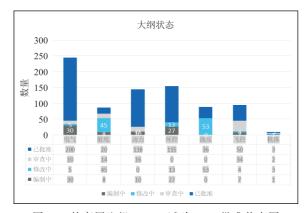


图 16 某商用飞机 MOC9 试验 QTP 批准状态图 Fig. 16 XX commercial aircraft MOC9 test QTP approval status

## 5 结 论

1)本文对商用飞机 MOC9试验项目应用大数据分析方法进行可视化展示开展研究,在项目策划阶段进行数据仓库设计,建立多层次多维度的可视化模型,帮助项目经理在掌握项目整体情况的同时,深入挖掘各设备每项要素的研制风险和

问题。

2)本文研究结果应用到MOC9试验项目中把工作分解结构(WBS)中各要素的研制历程通过建立量化的可视化模型能更准确直观地判断项目团队各专业、各供应商和局方的效率,挖掘出主制造商和供应商人力资源是否充足、技术能力是否达到要求、管理流程设计是否合理,并帮助项目经理快速精准找到问题的根本原因或者识别风险,提前做出预警,实现风险自动化识别预警,为管理人员提供更多便捷。

#### 参考文献

- [1] 汉斯一亨利奇·阿尔特菲尔德. 商用飞机项目[M]. 北京: 航空工业出版社,2014: 48-49. ALTFIELD H H. Commercial aircraft projects: managing
  - ALTFIELD H H. Commercial aircraft projects: managing the development of highly complex products [M]. Beijing: Aviation Industry Press, 2014: 48–49. (in Chinese)
- [2] 中国民用航空局航空器适航审定司. 航空器型号合格审定程序: AP-21-AA-2011-03-R4[S]. 北京: 中国民用航空局航空器适航审定司, 2011.
  - Department of Aircraft Airworthiness Certification of Civil Aviation Administration of China. Aircraft type certification procedure: AP-21-AA-2011-03-R4[S]. Beijing: Department of Aircraft Airworthiness Certification of Civil Aviation Administration of China, 2011. (in Chinese)
- [3] 谢春稳,周新铭,史新芳. 民机机载设备合格鉴定试验更改控制研究[J]. 工程与试验,2021,61(2):90-93.

  XIE Chunwen, ZHOU Xinming, SHI Xinfang. Study on change control of qualification test of civil aircraft airborne equipment[J]. Engineering & Test, 2021,61(2):90-93.

  (in Chinese)
- [4] 章光灿.民用飞机设备合格鉴定试验(MC9)研究[J].民用飞机设计与研究, 2020(1): 24-28.

  ZHANG Guangcan. Qualification test for civil aircraft equipment[J]. Civil Aircraft Design & Research, 2020(1): 24-28. (in Chinese)
- [5] LI Sujie, YANG Yi, YANG Lu, et al. Civil aircraft big data platform [C] // 2017 IEEE International Conference on Semantic Computing. San Diego, California, USA: IEEE, 2017; 328–333.
- [6] VIKAS D, CHITTARANJAN P. Project monitoring system for big data[C]// 2017 IEEE Region 10 Conference. Indore, India: IEEE, 2016: 1-5.
- [7] KUSEK G, KILIC I. Project-based application on big data usage [C] // Fourth International Conference on Agro-geoinformatics. US: IEEE, 2015: 89-92.
- [8] 魏壮宇,蔡红霞.大数据下基于关联规则算法的民机质量数据分析可视化[J]. 计量与测试技术,2018,45(4):66-68.

- WEI Zhuangyu, CAI Hongxia. Visualization of civil aircraft quality data analysis based on association rule algorithm under big data[J]. Measurement and Test Technology, 2018, 45(4): 66-68. (in Chinese)
- [9] 陈金,党帅,吴波.民机运行大数据分析平台整体架构研究[J]. 计算机测量与控制,2018,26(1):281-283,288. CHEN Jin, DANG Shuai, WU Bo. Research on big data analysis platform architecture of civil aircraft operation [J]. Computer Measure & Control, 2018, 26(1):281-283,288. (in Chinese)
- [10] 袁博,郝澜宇,陈震,等.基于文本挖掘的飞机装配质量问题智能诊断——以中国商飞公司为案例[J].工业工程与管理,2021(5):195-201.
  YUAN Bo, HAO Lanyu, CHEN Zhen, et al. Intelligent diagnosis of quality problem of aircraft assembly based on text mining: the case of COMAC[J]. Industrial Engineering

and Management, 2021(5): 195-201. (in Chinese)

- [11] 张磊磊,杨锋,刘贡平,等. 基于大数据的飞机试验业务综合管控架构规划与研究[C]//第十九届中国航空测控技术年会. 西安: 航空工业测控技术发展中心, 2022: 74-79. ZHANG Leilei, YANG Feng, LIU Gongping, et al. Planning and research on experimental task synthetical controlling framework of aircraft based on big data[C]// The 19th China Aviation Measure and Control Technology Annual Conference. Xi'an: Aerospace Industry Measurement and Control Technology Development Center, 2022: 74-79. (in China)
- [12] 赵学武,吴宁,王军,等. 航空大数据研究综述[J]. 计算机科学与探索,2021,15(6):1000-1001.

- ZHAO Xuewu, WU Ning, WANG Jun, et al. Overview of aviation big data research[J]. Journal of Frontiers of Computer Science and Technology, 2021, 15(6): 1000-1001. (in Chinese)
- [13] MANYIKA J, CHUI M, BROWN J, et al. Big data: the next frontier for innovation, competition and productivity [R]. US: McKinsey Global Institute, 2011.
- [14] 李军. 大数据——从海量到精准[M]. 北京: 清华大学出版社, 2014.
  - LI Jun. Big data: from massive to precise[M]. Beijing: Tsinghua University Press, 2014. (in Chinese)
- [15] RAMESH S, DURSUN D, EFRAIN T. 商务智能——数据分析的管理视角[M]. 北京: 机械工业出版社, 2018: 83-85.
  - RAMESH S, DURSUN D, EFRAIN T. Business intelligence: management perspective of data analysis [M]. Beijing: Machinery Industry Press, 2018: 83-85. (in Chinese)
- [16] ROBERT L. Python 数据挖掘入门与实践[M]. 北京:人 民邮电出版社, 2016: 81-102. ROBERT L. Python introduction and practice of data mining [M]. Beijing: People's Post and Telecommunications Press, 2016: 81-102. (in Chinese)
- [17] 姜维. 文本分析与文本挖掘[M]. 北京: 科学出版社, 2018: 93-96.
  - JIANG Wei. Text analysis and text mining [M]. Beijing: Science Press, 2018: 93-96.

(编辑:丛艳娟)