

T/CESA 1159-2021软件成熟度标准的软件验证和确认与实践

徐俊^{1,2}, 毛明志^{1*}, 李军³

¹中山大学, 广东 广州

²广州赛宝认证中心服务有限公司, 广东 广州

³广东省科技基础条件平台中心, 广东 广州

收稿日期: 2022年6月22日; 录用日期: 2022年8月15日; 发布日期: 2022年8月24日

摘要

本文分析了软件产品质量问题与改进难点, 对目前国内外主流的验证和确认成熟度模型实践条款进行了比对, 结合T/CESA 1159-2021软件成熟度标准软件测试过程要求, 分析了软件验证和确认与实践。

关键词

软件成熟度, 验证和确认(V&V), 软件测试

Software Verification and Validation and Practice of T/CESA 1159-2021 Software Maturity Standard

Jun Xu^{1,2}, Mingzhi Mao^{1*}, Jun Li³

¹Sun Yat-sen University, Guangzhou Guangdong

²Guangzhou CEPREI Certification Center, Guangzhou Guangdong

³Guangdong Science and Technology Basic Condition Platform Center, Guangzhou Guangdong

Received: Jun. 22nd, 2022; accepted: Aug. 15th, 2022; published: Aug. 24th, 2022

Abstract

This paper analyzes the quality problems and improvement difficulties of software products,

*通讯作者。

compares the current mainstream practical terms of verification and validation maturity model at home and abroad, and analyzes the software verification and validation and practice in combination with the requirements of software testing process of T/CESA 1159-2021 software maturity standard.

Keywords

Software Maturity, Verification and Validation (V&V), Software Test

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

验证和确认(Verification and Validation, V&V)是复杂工程系统可靠性认证中建模和模拟(Modeling and Simulation, M&S)置信度评估的重要手段[1]。软件验证和确认一直是计算机开发周期中最重要的阶段,目前,关于验证与确认的能力评价标准成为业界研究热门,在数值科学技术程序、应用软件界面设计、复杂工程建模和模拟、自主研发平台等方面,国内已有不少的研究[2] [3] [4]。然而,关于软件验证和确认的能力成熟度国产生态标准明显赶不上快速发展的软件行业发展需求。

2. 软件产品质量问题与改进难点

软件行业是信息产业的核心,提升关键软件技术创新和产业供给能力,对深化供给侧结构性改革,工业赋能以及促进新旧动能转换,推动互联网、大数据、人工智能和实体经济深度融合,建设现代化的经济体系具有重要意义。为了促进软件产业发展,2000年以后我国陆续出台了多项法规和政策,新一代信息技术提供了全新架构,给产业“换道超车”提供了重大历史机遇;同时开源软件技术发展也为软件产业发展提供了技术来源,降低了技术门槛和开发成本。我国在目前正在积极创造良好营商环境鼓励海外人才引进,人才引进进而带动技术引进,这些都给整个软件行业的发展带来了良好的契机。

但是目前整个软件行业发展也面临巨大挑战,我国软件产业大而不强,国内软件能力评价标准大幅度缺失,目前国外标准模型占据了市场主体,国内已有的个别标准模型评价体系明显陈旧且应用程度不高,现有的标准对软件产业引领作用发挥有限;软件行业自主生态薄弱,自主软件产业生态尚未建立;国内软件基础能力不足,核心关键技术领域技术尚未突破;产业价值失衡,软件作为硬件附属品观念没有根本扭转。发展环境有待完善:知识产权保护力度需进一步加强。软件行业还存在高端人才结构性短缺,公共服务能力不足等短板问题。

3. 国内外软件验证和确认相关成熟度标准情况

软件能力成熟度模型(CMMI2.0)——美国软件能力成熟度模型(CMMI)从1991年SEI发布的《CMM》开始发展至今,已经成为了目前全球最流行、最具影响力的软件生产过程标准,2018年CMMI Institute发布了最新的CMMI2.0模型[5]。CMMI评估目前占据了我国软件成熟度认证市场的95%以上市场份额。最近三年每年评价约3000家,证书三年有效,目前实施CMMI3级以上认证的企业数量约9000家以上。以2019数据为例,CMMI2级企业6家,占0.25%;CMMI3级2035家占84.65%;CMMI4级22家占0.92%;CMMI5级341家占14.18%。2019年中国大陆通过CMMI认证的企业数量共2403家企业,相比2018年

数量(1894家)增加了509家企业。

ISO/IEC15504《软件过程改进和能力测定》——ISO/IEC15504系列标准推出以来,在欧洲的应用反馈并未达到预期。目前在WG10组承担标准研制工作,标准编号转为ISO/IEC330XX系列。ISO/IEC相关33061为软件生存周期过程的过程评估模型[6]。

软件过程能力评估和软件能力成熟度评估体系(SPCA)——信息产业部会同国家认证认可监督管理委员会于2001年建立了软件过程能力评估和软件能力成熟度评估体系(SPCA)。两项行标发布至今,之前在国内应用的数量并不多,这两年由于信创工程的推进,评估数量快速上升,目前约占全国5%左右的市场份额,2021年全国SPCA评估企业数据约100家左右[7]。

GJB 5000B-2021《军用软件研制能力成熟度模型》——GJB 5000B-2021自2022年3月1日起正式实施规定了军用软件研制单位应达到的军用软件研制能力等级要求。军事代表机构负责受理软件研制能力评价申请,是强制性认证。不适用于商业化软件的评价与推广[8]。

Automotive SPICE 3.0 (Automotive Software process improvement and capability determination)——汽车软件过程改进和能力测定。它是车载软件开发的过程标准,用于欧洲整车厂对供应商进行软件过程评估。欧洲的主要汽车制造商共同策定的面向汽车行业的流程评估模型,意在改善和评估汽车电子控制器的系统和软件开发质量。按照汽车业SPICE的要求来审查其供应商的电子产品的开发的开发质量。目前国内个别汽车研制企业在进行评估工作,随着新能源汽车互联网汽车的快速发展,Automotive SPICE认证将在国内有很好的发展前景[9]。

T/CESA 1159-2021标准于2021年6月发布。该团体标准界定了软件能力成熟度的内容框架,主要包括实践域的分类、实践域和成熟度等级的定义,规定了组织治理、软件开发、项目管理和支持保障4个管理域的能力要求。规定了软件能力成熟度模型在不同等级中的实践活动要求。该标准主要适用于以下场景企业:寻求软件开发提供商并要求确保软件开发质量的顾客;希望展现其软件开发和交付能力成熟度的组织;通过标准的有效实施与运行来持续改进软件开发和交付绩效的组织;依据标准的要求实施符合性评价的第二方和第三方;符合性评价人员培训或建议的提供者[10]。

4. T/CESA 1159-2021 标准与 CMMI2.0/GJB5000B/SPCA 的验证确认过程对比

论文“验证”的目的在于保证工作产品满足其规定的要求。“验证”过程方面强调验证准备、验证执行和确定纠正措施。“验证”过程包括按照需求(包括顾客需求、产品需求和产品构件需求)对产品 and 中间产品进行验证。“验证”过程注定是一种渐进的过程,因为它要在产品和工作产品整个开发过程中执行,即,从对需求进行验证开始,然后是对推进中的工作产品进行验证,最后是对完成的产品进行验证。在产品每个层次上对工作产品的验证有助于提高产品满足顾客、产品和产品构件需求的可能性。

“确认”的目的在于证明,产品或产品构件当被置于其预定环境中时适合于其预定用途。

“确认”过程要证明所建造的产品将在其预定环境中发挥其预定作用。各项确认活动的做法。

与验证类似(例如,测试、分析、仿真等等)。确认活动和验证活动往往同时进行,并且可能利用同一个环境的某些部分。其间差别在于验证是证明产品符合产品规格说明的要求,而确认是证明产品适合于在预定运行环境中使用。换句话说,验证保证“做得正确”,而确认则保证“做的东西正确”。

如果可能,应该采用将在其预定环境运行的实际产品进行确认。可以使用整个环境,也可以使用一部分。通过早期开展确认活动(例如对照顾客和最终用户的运行需要对顾客需求进行确认),可以在开发生存周期的早期发现问题。“验证”过程方面与“确认”过程方面看起来类似,但是它们处理的问题不同。

“确认”是要证明所提供的(或将要提供的)产品适合其预计的用途,而“验证”则是要查明工作产品是否恰当地反映了规定的要求。换句话说,验证要保证“做得正确”,而确认则要保证“做的东西正确”。

验证与确认的比对见表 1。

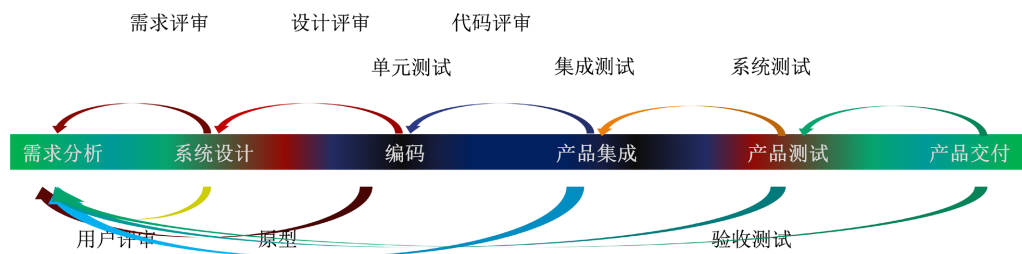
Table 1. Comparison table of verification and validation

表 1. 验证与确认的对比表

属性	验证	确认
定义	确保所选择的工作产品满足特定的需求	确保产品或产品构件能够在预期的环境下使用
重点	做法是否正确，强调做事的过程。 “正确的做事”	结果是否正确，强调做事的结果。 “做正确的事”
参照物	上一阶段的工作产品	原始需求
可采用的方法	评审、测试、审查	评审、测试、审查、原型

验证与确认活动示例见图 1。

验证活动示例



确认活动示例

Figure 1. Examples of verification and validation activities

图 1. 验证与确认活动示例

软件测试是证实软件部件或软件系统的需求是否完成且正确，每一阶段工作软件是否实现在上一阶段规定的需求或条件，以及最后的软件是否满足指定的需求的过程。

T/CESA 1159-2021 标准软件测试过程目标是为了验证软件是否满足指定的需求；确认软件在预定环境下是否实现其预期用途。

T/CESA 1159-2021 测试过程成熟度二级实践包括如下：

实践 1：建立软件系统测试方案。组织应建立软件系统的测试方案，确定软件系统测试的范围、职责、使用的工具和方法、进度和周期、规程和准则、测试资源、测试报告文档要求等。

实践 2：测试准备。测试人员为完成测试活动，应建立、维护测试用例，并准备测试环境和必要的工具。

实践 3：执行测试，记录结果。依据测试方案执行测试，记录测试执行的结果，对测试过程进行总结形成测试报告。

成熟度三级实践包括如下：

实践 1：建立和使用组织测试过程资产。

实践 2：分析测试数据。应从组织级和项目级对测试数据进行综合分析，识别改进测试过程的机会。

验证与确认在不同模型体系的区别见如下表 2。

Table 2. Comparison table of differences between validation and validation in different model systems

表 2. 验证与确认在不同模型体系的区别对照表

等级	T/CESA 1159-2021 测试过程	CMMI2.0 验证与确认	GJB5000B/SJT11235 验证/确认过程
一级实践	无	1.1 进行验证 1.2 进行确认。	无
二级实践	2.1 建立软件系统 测试方案 2.2 测试准备。 2.3 执行测试， 记录结果	2.1 为验证和确认选择组件和方法。 2.2 开发、保持更新和使用支持验证 和确认所需的环境。 2.3 开发、保持更新和遵循验证和 确认的规程。	无
三级实践	3.1 建立和使用组织 测试过程资产 3.2 分析测试数据	3.1 开发、保持更新和使用验证和 确认的准则。 3.2 分析和沟通验证和确认结果。	SG 1 准备验证/确认 SP 1.1 建立验证/确认策略 SP 1.2 建立验证/确认环境 SP 1.3 建立详细验证/确认计划 SG 2 执行同行审查 SP 2.1 准备同行审查 SP 2.2 进行同行审查 SP 2.3 分析同行审查数据 SG 3 验证/确认所选择的工作产品 SP 3.1 进行验证/确认 SP 3.2 分析验证/确认结果和确定纠正措施 SP 3.3 进行复验证

四个标准实践的相同点：都包括建立软件系统测试方案，做好测试准备。执行测试，记录结果，分析测试数据。

四个标准实践的不同点如下：

1) T/CESA 1159-2021 由于 1 级为无序管理级，有部分的软件交付活动，工作完成结果无法确定，软件项目交付成果可能成功，故在 1 级没有过程定义。CMMI2.0 的 0 级为不完整级，无计划且未知，可能可以或无法完成工作。1 级是初步级，不可预测且被动。可以完成工作，但通常会延期且超预算，故在 1 级有基本的工作要求。SJT11235-软件能力成熟度模型和 GJB5000B 属于阶段式迭加要求，都是参考 CMMI1.0 标准演化过来的，验证与确认过程属于三级已定义级，故在 1 级和 2 级没有验证确认的相关实践要求。

2) T/CESA 1159-2021 和 CMMI2.0 把同行评审从验证确认过程中单列出去，建立了单独的同行评审过程，SJT11235 和 GJB5000B 是参照早期 CMMI1.1 标准，验证过程中包括同行评审相关实践要求。

3) T/CESA 1159-2021 测试过程 3.1 建立和使用组织测试过程资产，是其他两个标准过程所没有的。标准要求组织应具备较完备测试运行管理机制，包括文件化的测试规程和准则、相关职责、测试工具、环境资源等，配备有独立的系统测试团队，各项测试活动按计划协调开展。为提高测试的效率和质量，建立组织测试过程资产规范，将可供后续工作重用的测试方案、测试用例、仿真模块、驱动、数据等进行提炼，形成组织测试过程资产。CMMI2.0 的 VV 过程 2.3 要求提及开发、保持更新和遵循验证和确认的规程，只是局限在建立验证和确认程序。

5. T/CESA 1159-2021 标准软件验证和确认工作实践

常见的测试活动的主要覆盖内容如下表 3 所示。

Table 3. Table of coverage of test activities

表 3. 测试活动的覆盖内容表

名称	测试对象	侧重点	参照物	充分性的评价方法	时机	测试方法	测试执行者
单元测试	软件的最小单元,如函数、方法等	逻辑的正确性	详细设计、源程序	代码、分支等覆盖率	软件中的基本组成单位完成后,边开发边测试	白盒测试、动态测试	一般是开发人员
集成测试	软件的模块、子系统	接口的正确性	概要设计、详细设计	接口覆盖率	软件系统集成过程中,边集成,边测试	黑盒测试、功能测试、白盒测试等	开发人员与测试人员
系统测试	系统	需求的满足性	产品需求	用户场景覆盖率	系统开发完成后,交付客户之前	黑盒测试、功能测试、非功能测试等	测试人员
验收测试	系统	需求的满足性	客户需求	需求覆盖率	交付客户后,正式投入使用之前	黑盒测试、功能测试、非功能测试等	客户

5.1. 验证和确认的工作产品

验证活动工作产品选择包括单元测试对象选择,单元模块功能测试可以选择核心代码、多循环、多判断层次的复杂度高的代码。代码走查轮查评审可以选择新手的代码、核心代码执行评审。软件开发过程文档同行评审,可以选择需求、设计等工程类文档。确认工作产品选择包括通常在系统测试计划、验收测试计划或者试运行计划中标识确认的工作产品。

5.2. 选择验证和确认的方法

验证的常用方法包括需求检查、演示、负荷、压力和性能测试、功能、接口或连接,以及集成测试、原型制造、建模和仿真。确认方法包括与最终用户一起审查、原型展示、功能展示、试运行、由最终用户和其他受影响的干系人进行的解决方案组件测试等。

5.3. 建立维护支持验证和确认的环境

单元、集成、系统、验收测试的环境需要提前建立,例如:测试的工具、服务器等要提前配置好。评审环境会议室、投影仪、提前确认参与评审的工作产品、时间、地点、评审方式、评审通过原则等。

5.4. 为选择的工作产品建立和维护验证和确认的规程和准则

测试规程包括测试流程(单元、集成、系统、验收),测试规范,通过准则。通常需要研发组织定义单元测试、集成测试、系统测试、验收测试规程,项目可以进行裁剪。需要项目制定量化的通过标准,如:严重 BUG 全部解决,遗留的缺陷密度不能多于一个具体数值,比如小于 0.1 个 BUG/KLOC。

5.5. 分析验证和确认活动的结果

分析测试用例的执行结果,将实际测试结果和预期结果对比,同时记录缺陷并分析、解决缺陷;建立单元、集成、系统/验收测试报告,记录验证和确认结论。执行测试结果分析,确认测试结果是否达到通过,修改后通过或者不通过,重新修改等。举例来说可以在测试报告中进行分析,比如针对 BUG

类型分类分析, 确认 bug 类型是功能、性能、界面、数据; bug 严重等级; BUG 移除, BUG 引入阶段 BUG 模块分布; BUG 趋势分布。组织建立 BUG 数据库, 分析 BUG, 建立 BUG 预警机制。

6. 结论

目前主流的三个民用成熟度标准模型都对验证和确认过程进行了实践要求, 由于成熟度模型更多是过程改进的框架性要求, 无法对相关的实践要求进行进一步细化标准条款。我们需要在模型或标准的指导下, 进一步结合不同行业软件的研发特点进行细化验证和确认相关的条款实施细节要求, 由于这几年国内信创工程的推出, 中美关系的不确定因素等国内外环境因素变化, 建立适合于中国的国情以及中国软件企业的特点的国产化标准工作凸显重要地位, 今后需要进一步加强验证和确认相关的国产化标准的研制和细化。后续需要在 T/CESA 1159-2021 基础上聚焦验证确认核心能力, 建立不同行业组织结合自身软件在合理范围内对标准内容进行裁剪通用标准, 进一步优化软件改进生态, 推动我国软件行业高质量发展。

基金项目

本文得到中山大学自选课题的资助。

参考文献

- [1] 王瑞利, 温万治. 复杂工程建模和模拟的验证与确认[J]. 计算机辅助工程, 2014, 23(4): 61-68.
- [2] 王瑞利, 林忠, 袁国兴. 科学计算程序的验证和确认[J]. 北京理工大学学报, 2010, 30(3): 353-356+360.
- [3] 王瑞利, 梁霄, 喻强, 林忠, 林文洲. 自主研发软件 LAD2D 的验证、确认及不确定度量策略[C]//第十二届中国 CAE 工程分析技术年会论文集: 2016 年卷. 2016: 31-39+17.
- [4] 毛明志, 李建国. 用户界面设计及对我国数值软件设计的启示[C]//第三届全国计算爆炸力学会议论文集《计算爆炸力学进展》: 2006 年卷. 2006: 11-15.
- [5] CMMI Institute. 《软件能力成熟度模型(CMMI2.0)》[S]. 2008.
- [6] ISO&IEC. ISO/IEC15504 TR 《软件过程改进和能力测定》[S]. 1998.
- [7] 信息产业部. SJ/T 11235-2001 软件过程能力评估和软件能力成熟度评估体系(SPCA) [S]. 2001.
- [8] 中央军委装备发展部. GJB 5000B-2021 《军用软件研制能力成熟度模型》[S]. 2021.
- [9] HIS (2017) Automotive SPICE 3.1 (Automotive Software Process Improvement and Capability Determination).
- [10] 中国电子工业标准化技术协会. T/CESA 1159-2021 《软件能力成熟度模型》[S]. 2021.