

测试成熟度模型集成(TMMi®): 全球测试成熟度和认证趋势

Vahid Garousi
Queen's University Belfast
Bahar Software Engineering Consulting Limited

Erik van Veenendaal
TMMi Foundation

中文发布版1.0
由TMMi®基金会中国分会负责翻译

编辑:

英国贝尔法斯特女王大学, 巴哈软件工程咨询有限公司, Vahid Garousi
TMMi®基金会, Erik van Veenendaal

版权说明

受限于版权条款的不限制分发

版权归英国TMMi®基金会

参与中文本地化的TMMi®专家

主要内容	翻译专家	评审专家
全篇	冯云显	朱伟

摘要:

测试成熟度模型集成 (TMMi®) 是行业中用于软件测试实践的成熟度评估和能力改进的一种通用模型。最初受到能力成熟度模型集成 (CMMI®) 的启发, 并由 TMMi® 基金会管理, TMMi® 规格说明为评估和改进团队和组织的测试能力提供了详细的指导方针。在本文中, 我们在本文中介绍了 TMMi® 的现状报告、使用 TMMi® 的动机和好处、以及公司在各过程域的定级情况, 以及公司如何在其每个过程域中进行定级。

关键词:

测试成熟度模型集成, TMMi®, 趋势分析, 测试成熟度认证

1 介绍

为了满足对软件质量日益增长的需求, 自20世纪80年代以来, 软件行业中提出了各种新方案、模型和方法。例如能力成熟度模型集成 (CMMI®) 模型 (cmmiinstitute.com) 和 ISO/IEC 15504 标准, 也称为软件过程改进和能力确定 (SPICE) 模型。

在过去的三十年中, CMMI® 采用逐渐增多, 主要是从事政府和国防项目的组织。根据 CMMI® 研究所 (sas.cmiiinstitute.com/pars/pars.aspx) 的数据, 截至本文撰写之时 (2020年12月), 已有9650家公司获得了 CMMI® 的评估 (认证)。

尽管有研究报告称, CMMI® 和 SPICE 等模型, 但它们主要关注的是对“整体”软件开发过程的过程改进。这样的过程改进模型通常不会为软件测试提供具体的改进建议。因此, 人们开发了各种模型来专门改进软件测试实践。IEEE Software [1] 2018年的一篇调查论文报告了58个测试成熟度和能力改进模型的目录, 例如, 测试成熟度模型集成 (TMMi®) (tmmi.org)。

本文的作者之一来自 TMMi® 基金会, 他在本文中简要报告了全球 TMMi® 评估和认证的发展趋势。我们注意到, TMMi® 模型自2012年以来就存在了, 因此本文的目标当然不是介绍 TMMi®, 而是介绍使用 TMMi® 的动机和好处, 以及公司如何在其每个过程域中进行定级。这些数据是以匿名方式从 TMMi® 基金会的内部评估和认证数据库中汇编和准备的。

我们首先简要介绍 TMMi®。

2 TMMi® 的概述

TMMi® 的起源可以追溯到1988年发表的 Gelperin 和 Hetzel 的演进测试模型 [2], 以及早期名为测试成熟度模型 (test Maturity model, TMM) 的测试改进模型 [3]。由于看到了对更集中的测试改进模型的需求, 一些测试和质量专家 (志愿者) 于2005年聚集在一起 (主要在欧洲), 并成立了 TMMi® 基金会。TMMi® 定级的第一个版本 (1.0) [4], 由基金会于2012年提出。在撰写本文时, 定级的最新版本为 1.2 [5] (于2018年发布)。

TMMi® 框架 [5] 是测试过程改进的指南和参考框架。TMMi® 使用成熟度级别的概念进行过程评估和改进。此外, 对于每个成熟度级别, 都确定了一组过程域、目标和实践。该基金会能够得到 TMMi® 地方分会提供支持, 这些分会在其所在国家/地区提供和组织基于 TMMi® 的咨询服务。在撰写本文时, TMMi® 已在中国、美国、西班牙、巴西和法国等地建立了22个地方分会。

TMMi® 与国际测试标准以及国际软件测试资格委员会 (ISTQB®) 的教学大纲和术语保持一致。TMMi® 基金会会有意识地不引入新的或他们自己的术语, 而是重用 ISTQB® 术语。这对所有获得 ISTQB® 认证的测试专业人员来说都是一个优势 (在撰写本文时, 全球大约有700,000名测试专业人员)。有了 TMMi®, 企业可以让经过认证的评估人员对其测试流程进行客观评估。测试专业人员和顾问个人也可以获得“TMMi® 专业人员”认证。

测试专业人员和顾问也可以通过个人认证成为“TMMi® 专业人员”。

与 CMMI® 类似, TMMi® 也有一个“分阶段”的测试流程评估和改进方案。它包含了一个组织在其测试过程从临时的, 也称为“初始级或未管理级” (级别=1) 发展到已管理级 (级别=2)、已定义级 (级别=3)、已测量级 (级别=4) 和优化级 (级别=5) 时所经过的阶段或级别。

图1显示了 TMMi® 的五个成熟度级别, 它们的16个过程域 (PA), 以及作为元模型的 TMMi® 结构。测试团队/组织要达到任何级别, 需要满足该级别和较低级别的所有过程域。

在分层设置中, 每个 PA 都有几个特定目标 (SG)、特定实践 (SP)、子实践 (SP)、通用目标 (GG) 和通

用实践（GP）。在这五个级别中，总共有50个SG、173个SP、845个GP、32个GG和192个GP。有关这些要素的详细信息请参见TMMi定级[5]中找到。

(Define test goals), and SP 1.2 (Define test policy).

例如，成熟度级别2（“已管理级”）有五个PA，例如PA 2.1（测试方针与策略）。该PA有三个SG：SG1（建立测试方针）、SG2（建立测试策略）和SG3（建立测试绩效指标）。上述SG1依次具有三个SP，例如SP1.1（定义测试目标）和SP1.2（定义测试方针）。

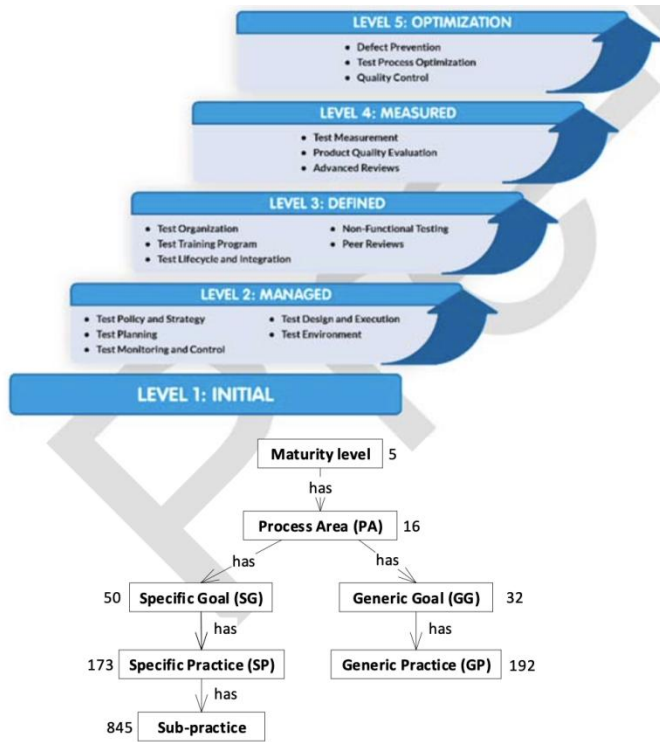


图1-顶部：TMMi®及其过程域(PA)的成熟度级别。底部：作为元模型的TMMi®结构。

TMMi®的一个主要基本原理：它是一个适用于各种生命周期模型的“通用”模型。该模型已被翻译成多种语言，如西班牙语、法语、葡萄牙语和中文。TMMi®行业应用的一些经验报告和案例研究已经发表，例如[1, 6-8]。

例如，一项研究[7]报道了爱沙尼亚一家游戏软件公司使用TMMi®改进测试过程的单对象案例研究。该研究实证发现，TMMi®框架在参考模型、评估方法和数据提交要求方面的全面而详细的文件为评估和改进组织提供了充分的支持。

在另一项研究[8]中，TMMi®和TPI-Next都应用于一家瑞典大型公司（VolvoIT）。研究发现，尽管这两个模型整体表现出很强的相似性，但由于它们的模型表示不同，评估结果的差异是显而易见的。评估结果的映射和比较表明，TMMi®对成熟度水平的要求比TPI-Next严格得多，更难达到。此外，“TMMi®中的工作示例对测试过程进行了非常详细的描述，这为进行评估提供了很好的指导”。此外，行业研究发现，为了

成功应用这两种方法，软件测试的扩展知识至关重要。

A logical question to explore is about the motivations

需要探讨的一个合乎逻辑的问题是，公司使用TMMi®评估和改进流程的动机。涉及整个软件开发生命周期（SDLC）的成熟度模型，例如CMMI®，仅为软件测试提供一些高阶评估和改进标准，因此，对测试过程改进（TPI）的使用有限。通过调查学术文献和非公开出版文献，一项研究[1]报告称，总体而言，使用TPI模型的三个主要驱动因素是：提高软件质量的需求、降低测试成本的需求以及过程和操作需求。TMMi®基金会最近在2020年进行的一项调查（结果尚未公布）显示，采用TMMi®的主要原因（动机）是：提高产品质量，降低产品风险，提高测试生产力（效率），与国际通用模型进行比较，以及提高测试团队的声望。

如前所述，TPI有超过58个模型[1]。我们有兴趣客观地比较不同模型的行业传播/渗透情况。然而，这方面的数据非常有限。在一项调查研究[1]中，回顾了来自行业 and 学术界的181个来源，181个来源中有57个使用了TMMi®，而18个使用了名为TPI-Next[9]的模型。由于缺乏关于成熟度模型在业界受欢迎程度的客观统计数据，上述数字可以作为TMMi®和TPI-Next的受欢迎程度的指标。

TMMi®是一个独立的模型，由TMMi®基金会管理，但TPI-Next模型是由一家软件咨询公司开发的。我们在表1中比较了TMMi®和TPI-Next。

表1-TMMi®和TPI-Next比较

准则	TMMi®	TPI-Next
表达方式	阶段式模型	连续式模型
测试级别	所有测试级别（单元、集成、系统、用户体验测试）	聚焦在高阶测试级别（系统和用户体验测试）
支持的测试方法	独立的测试方法	关联到TMap (tmap.net)
术语	以ISTQB为基础	以TMap为基础
基于的软件过程改进模型	关联到CMMI	无
认证	可以通过正式的评估进行认证	无

关于TMMi®与其他现有模型比较，早先的一项研究[1]指出了TMMi®和TPI-Next都是用于测试过程改进的“通用”模型，还有一些“专用”模型，如单元测试成熟度模型（UTMM）和自动化软件测试成熟度模型（ASTMM）[1]。从这一领域的58个模型[1]来看，从评估的完整性角度来看，确定最“有希望”的模型并不容易。特定的团队/公司会根据自身需求，选择并应用合适的改进模型。

3 不断增长的认证数量

根据TMMi®基金会的内部认证数据库，截至2019年底，全球已有187家公司向基金会提交了评估申请。73家公

司申请了“非正式”评估，因为他们的目的是获得“指示性[评级]，而不是正式的成熟度评级”[5]。114家公司申请了“正式”评估，因为它们需要接受“详细的评估和/或考虑走上认证之路”。其中，已经颁发了111份证书（名单详见tmmi.org）。所有评估均由TMMi®认证（首席）评估员团队系统地进行。对于评估，除了规范[5]之外，还有另一份名为TMMi®评估方法应用要求（TAMAR）[10]的文件，用于确保评估的严谨性和一致性。

图2显示了2011-2019年间TMMi®评估的累积趋势，从中可以看出人们对获得TMMi®评估的兴趣越来越大。

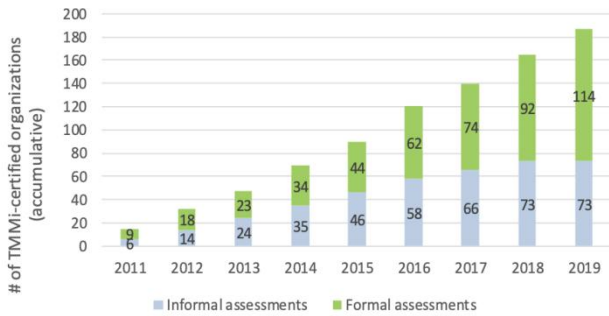


图2-自2011年以来，TMMi®评估和认证数量不断增加

在分析评估和认证数据库时，我们希望深入了解TMMi®评估中代表性最高和最低的地区和/国家。在图3中，我们展示了这些数据。英国、韩国和中国是排名前三的国家。有趣的是，某些国家的参与率更高。例如，在英国，TMMi®已被政府部门广泛采用，也有媒体报道。这在一定程度上是因为基金会在某些国家加大了宣传力度（由于其成员和地方分会的努力）。我们应该补充一点，基金会最初是从英国开始的。

此外，我们认为某些国家的工作文化（“流程导向”），以及CMMI®在某些地区的大量存在，是影响TMMi®在某些国家更高普及率的其他因素。

请注意，图3中仍然缺少一些重要的IT活跃国家/地区。预计这种情况很快就会改变。例如，最近在加拿大、意大利和土耳其成立了三个TMMi®地方分会。德国和俄罗斯等其他国家也正在筹建TMMi®地方分会。

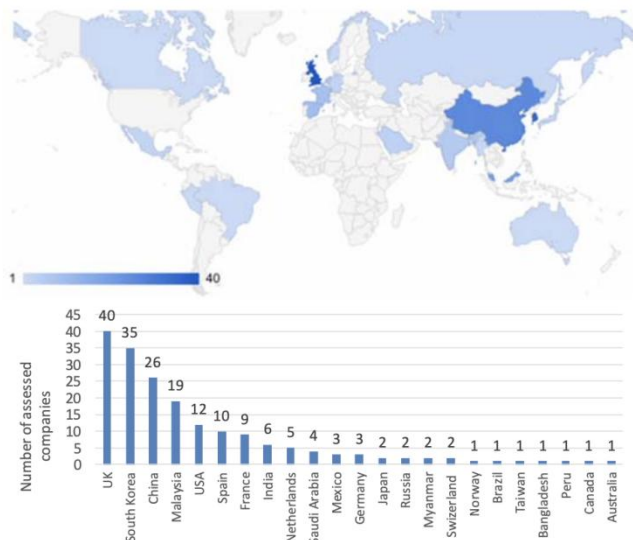


图3-接受TMMi®评估的组织国家(地区)

更有意义地了解上述国家数据的一种方法是将其与基金会数据库中的另一个指标，即经认证的TMMi®专业人员的所在地（居住国）相关联。TMMi® Professional是一个向所有对TMMi®感兴趣的人开放的认证。持有TMMi®专业证书是成为经认可的TMMi®评估师的先决条件，持有TMMi®评估师证书的人员有权进行TMMi®评估。

根据汇编的数据，我们在图4中展示了各国（截至2019年底）TMMi®专业人员数量与这些国家被评估组织数量的散点图。请注意，为了简洁起见，本图只显示了在两个指标方面都具有较高代表性的国家（每个指标至少有3个）。两个序列的皮尔逊相关系数为0.51，表示存在中等程度的相关性。因此，我们可以说，一般来说，一个国家的TMMi®专业人员数量越高，这个国家就更有可能有更多的接受评估的组织。这的确是意料之中的事，因为当一个国家有更多的TMMi®专业人员时，他们有助于提高对TMMi®的认识，也鼓励更多的组织申请评估和认证。在图4的散点图中，我们还可以看到其他各种有趣的现象，例如，印度拥有第三大TMMi®专业社区，但在TMMi®认证机构的数量排名却靠后。这可能是因为没有看到足够的动机/理由获得认证，例如，政府机构没有强制执行，如上文讨论的英国案例一样。此外，请注意，并非所有使用TMMi®的组织都会去申请认证。可能有许多专业人员在从事与TMMi®相关工作，但对于组织而言，“仅仅”获得利益（如在软件质量方面）往往就足够了，他们可能不会去申请认证

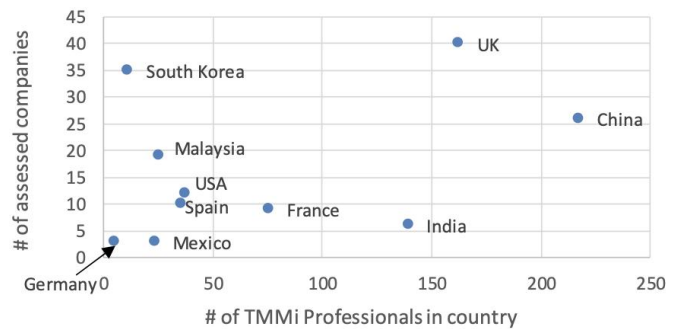


图4- TMMi®专业人员所在国家与被评估组织数量

4 公司达到的TMMi®级别

认证数据库中的其他数据是每个过程域（PA）的“等级”和每个申请公司或组织单元实现的具体目标（SG）。

我们在图5中展示了一个“单值”图，包括在所有认证（n=111）中达到的TMMi®级别的“移动”平均值（长度=10）。我们还研究了每年的中值和平均值。正如我们所看到的，主要是在八年的窗口期间移动平均值在3到4之间波动。例如，2011年，向TMMi®级别比较低（1-3级）的公司颁发了8份证书。总的来说，我们没有观察到这些年来在认证中获得的TMMi®级别有任何增长趋势。主要原因是，在某一年，任何成熟度级别的公司都可以申请并获得不同级别的认证。此外，成熟度级别数据集不适用于不同年份的同一组公司。



图5-每年通过认证达到的TMMi级别的平均值和中位数

5 每个过程域的分

我们还掌握了每个PA的详细数据。让我们从图1中回顾一下，在TMMi®的五个级别中，总共有16个PA。我们想知道，是否有一些特定的PA，让企业在通过这些PA时会遇到更多的“挑战”。值得一提的是，为了对给定申请人（公司或团队）的每个PA的分数进行排名，TMMi® TAMAR文件[10]提供了一个5分（量表）：

- (1) 完全实现（或完全实施），
- (2) 基本实现，
- (3) 部分实现，
- (4) 未实现，
- (5) 未审查；这与CMMI®评估类似。

图6显示了申请正式评估的114家公司的五级评分和他们16个PA的单独推栈图。正如我们所看到的，大多数申请者的PA都被评为“完全达标”。一些PA“基本上实现了”。部分实现或未实现占很小的比例。对于许多申请者来说，4级和5级的PA“没有经过审查”，因为在提交申请时，公司会指定他们打算进行评估的级别。大多数申请针对的是2级或3级，因此不必评估4级和5级的PA。

如果范围内的所有PA都被评为“基本达到”或“完全达到”，则组织可以将成熟度级别评为“达到”。此外，如果不能达到更低的成熟度级别，就无法达到更高的成熟度。因此，至关重要的是，特定的申请公司确保其有证据/能力“基本”或“完全”达到（满足）其打算获得认证的所有PA域。

回到我们的问题，即是否存在对公司来说通常更具挑战性的特定PA，我们发现与其他PA相比，PA2.5（测试环境）和PA3.4（非功能测试）的“未达到”分数相对更多（这在图6中可见）。这两个PA似乎在评估池中存在一定的挑战，例如，可能是一些公司在建立系统测试环境和进行非功能测试方面投入不足。非功能性通常是一种特定的专业知识，与其他过程域相比，它需要更多的资源投入，而且往往更难做好。因此，对于考虑申请的团队来说，一个可行的见解是确保改进团队中的这些方面。

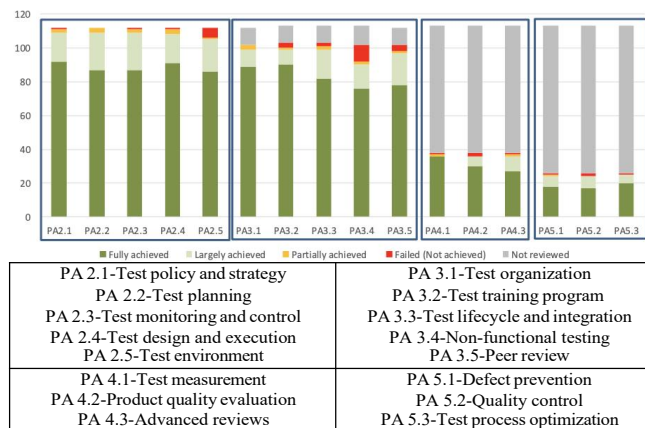


图6-接受正式评估的114家公司每个过程域(PA)的得分情况

6 TMMi®结束语和TMMi®的未来之路

这份简短的现状报告旨在深入了解全球TMMi®评估和认证的趋势，以及公司在TMMi®成熟度级别中等级的概况。

当然，还有一些与TMMi®相关的其他重要问题值得研究，例如TMMi®认证对组织开发的软件质量的影响。TMMi®基金会最近在2020年进行的一项调查的数据（结果尚未公布）可能会为这个重要问题提供一些参考。参与调查的三家TMMi®三级公司的自我报告缺陷检测百分比（DDP）分别提高了10%、20%和22%（平均17%）。我们要提醒的是，DDP是一个测试阶段发现的缺陷数量，除以该测试阶段和之后的任何其他方法发现的缺陷数量之和。因此，这些是TMMi®认证对软件团队的影响/好处的一些证据。

尽管在过去几十年中有许多软件质量举措，但软件行业仍在努力提供完美（无缺陷）的软件。很明显，要实现更高水平的产品质量，就需要更高级别的测试成熟度。随着越来越多的人意识到测试不当的代价，例如[11]，行业正在软件测试上投入更多的资源。公司正在寻找以更有效和高效的方式进行测试的方法，通常是通过改进测试流程。TMMi®是实现这一目标的既定手段之一。据报道，使用TMMi®对产品质量（测试效果）和测试效率[12]、减少测试执行时间和提高缺陷检测百分比（DDP）都有各种好处。

由于TMMi®和CMMI®之间有相似之处，因此对它们进行比较也很重要。从20世纪80年代末开始，CMMI®已经成为一种流行的软件过程改进模型，在世界范围内得到了广泛的应用，主要集中在政府和国防部门。事实上，CMMI®最初是美国国防部（DoD）要求采用的，这对普及起到了巨大的推动作用。就TMMi®而言，这种推动力正在慢慢形成。例如，马来西亚政府在2018年发布了一项政策，规定只有达到TMMi3级或以上的公司才能被指定为公共部门ICT项目的IV&V（独立验证和确认）供货商（bit.ly/MalaysiaTMMi）。

我们在本文中的分析表明，自2011年开始TMMi®评估以来，每年的年度评估数量在15-30家公司之间。2019年，TMMi®的使用量已经有所增长（如图2所示）。

TMMi®最近（自2017年以来）的另一个有趣的发展可能是建立了TMMi®地方分会。TMMi®地方分会可确保在当地提供TMMi®专业培训和评估服务。从最近的增长数字来看，这种做法已经取得了成功。

此外，TMMi®基金会和世界领先的测试认证组织ISTQB®于2019年年中结成联盟，共同进一步推动软件测试行业的发展。该联盟旨在将测试人员（ISTQB®）和过程（TMMi®）方面聚集在一起。

基于上述情况，我们预测，随着越来越多的公司看到TMMi®的好处，TMMi®在未来几年将得到更多的“青睐”。

References

- [1] Vahid Garousi, Michael Felderer, and Tuna Hacaloğlu, "What we know about software test maturity and test process improvement," *IEEE Software*, vol. 35, no. 1, pp. 84 - 92, 2018.
- [2] David Gelperin and Bill Hetzel, "The growth of software testing," *Communications of the ACM*, vol. 31, no. 6, pp. 687 - 695, 1988.
- [3] Ilene Burnstein, Ariya Homyen, Robert Grom, and C.R. Carlson, "A Model to Assess Testing Process Maturity," *Crosstalk: The Journal of Defense Software Engineering*, vol. 11, 1998.
- [4] TMMI Foundation, "TMMI specification (reference model), release 1.0," http://www.tmmi.org/pdf/TMMI_Framework.pdf, Last accessed: Oct. 2015.
- [5] TMMI Foundation, "Test Maturity Model integration (TMMi) reference model (specification), release 1.2," <https://www.tmmi.org/tmmi - documents/>, 2018, Last accessed: May 2020.
- [6] Erik van Veenendaal, Chaobo Shang, and Yihua Xu, "Achieving TMMi Level 3 - A Chinese Case Study," *Quality Matters magazine*, vol. 8, 2019.
- [7] Kerli Rungi and Raimundas Matulevičius, "Empirical Analysis of the Test Maturity Model Integration (TMMi)," in *Information and Software Technologies*, vol. 403, 2013, ch. 32, pp. 376 - 391.
- [8] Snehal Alone and Kerstin Glocksien, "Evaluation of Test Process Improvement approaches: An industrial case study," *Master of Science Thesis in Software Engineering, University of Gothenburg, Sweden*, 2013.
- [9] Alexander van Ewijk, Bert Linker, Marcel van Oosterwijk, and Ben Visser, *TPI next: business driven test process improvement*. Kleine Uil, 2013.
- [10] TMMI Foundation, "TMMi Assessment Method Application Requirements (TAMAR), Version 2.0,"
- [11] Tom Britton, Lisa Jeng, Graham Carver, Paul Cheak, and Tomer Katzenellenbogen, "Reversible Debugging Software," *University of Cambridge, Judge Business School, Technical Report*, 2013.
- [12] Erik van Veenendaal, "TMMi: The World Standard for Test Process Improvement," *Quality Matters magazine*, vol. 7, 2018.

作者简介

Vahid Garousi是Bahar软件工程咨询公司的经理，也是贝尔法斯特女王大学的副教授。他的专业知识包括

软件测试和改进软件工程实践。他在卡尔顿大学获得了软件工程博士学位。

欲了解更多信息，请访问：www.vgarousi.com。联系方式：v.garousi@qub.ac.uk

Erik van Veenendaal是TMMi基金会的首席执行官，曾任埃因霍温理工大学的高级讲师。他的专业知识包括软件测试和测试改进。他在蒂尔堡大学获得了商业经济学硕士学位。

联系方式：erik@erikvanveenendaal.nl