

协作软件助你事半功倍

“持续集成”服务既省时又保护关键代码

Sebastian Neubert 是德国海德堡大学的一名粒子物理学家,他带领一个团队研究名为“五夸克”的亚原子粒子。6名团队成员都可以访问用于运行其多步分析的软件代码,程序员每天都要更新其特性并进行漏洞修复。然而,随着每一个代码的变化,它们都有可能导入偶然的错误,从而扰乱基础算法。

为了防止这种情况发生,该团队对他们的分析进行了一次的核查,并使用错误检查算法,每当提出一项更改时,它们可以发挥功能,确保软件正常工作。例如,一项测试验证了当在实践数据上运行时,噪声消除算法能给出正确的输出。

2015年,为了节省时间和资源,Neubert 团队从技术产业中汲取灵感,用一种叫作“持续集成”的过程将的测试自动化。

持续集成

在持续集成中,对软件代码的更改会自动触发重复任务,例如错误检查。从根本上说,这个过程简化了程序员要完成的任务。程序员通常会写一份定期运行的测试清单,以确保他们的代码有效,正如 Neubert 团队所做的那样。但一个忙碌的团队可能会忘记或者没时间运行它们,从而导致出现错误。持续集成则让这一过程自动化,因此每当发生更改时,那些检查就会运行,从而节省了团队成员寻找错误所花费的时间。进行基因组分析的团队可以花更多的工作台上,而开发气候预测软件的小组则可以更好地改进模型。也就是说,由此带来的平和心态和测试本身一样好:一个设计糟糕的测试可以让错误在不被发现的情况下通过。

这一过程在商业和开源领域很常见。在新加坡举行的 2016 IEEE/ACM 国际自动化软件工程会议上报告的一项研究发现,在编码协作网站 GitHub 上托管的 34544 个最受欢迎的开源项目中,约有 40% 使用了持续集成。

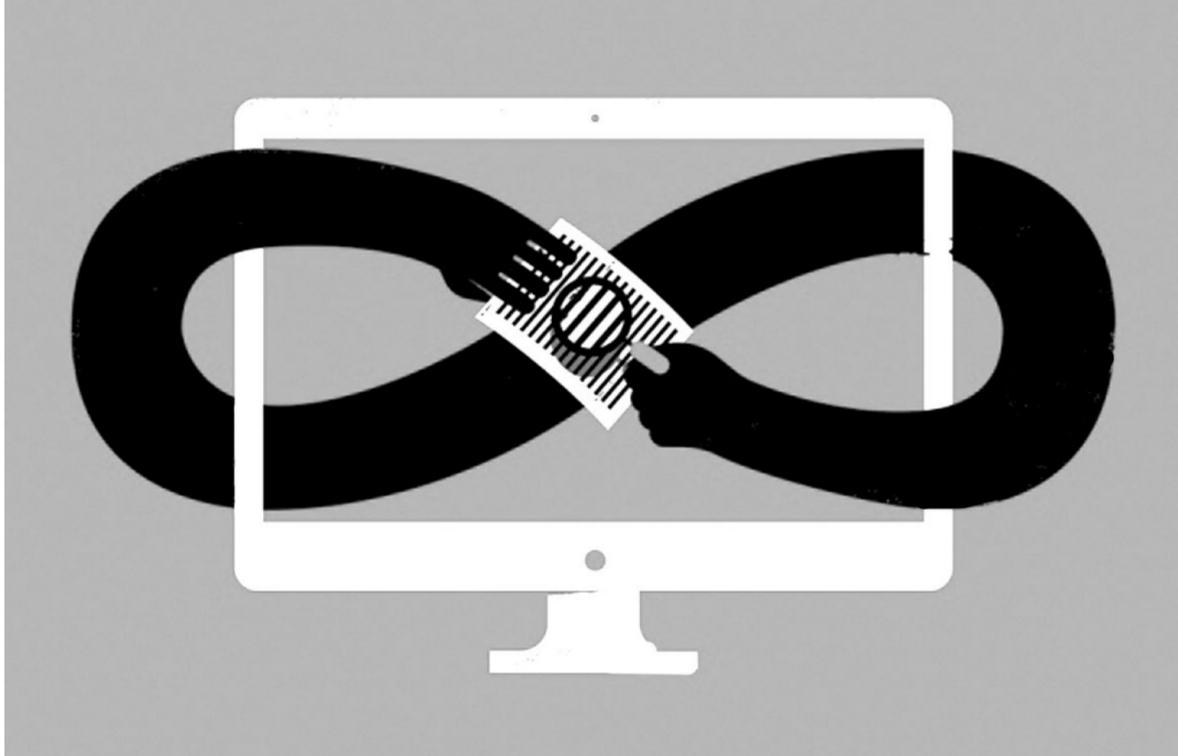
仅有少数几个开源项目或被认为是科学的软件,但越来越多的科学家正在寻求持续集成,从而让从测试代码到用最新数据更新文件等各种耗时的任务自动化。

瑞士日内瓦附近的欧洲粒子加速器实验室等类似机构的研究人员、位于美国华盛顿里奇兰太平洋西北国家实验室的研究人员以及加拿大多伦多安大略癌症研究所的研究人员已经接受了这种做法,但是其在科学领域的采用仍然相对稀少。

对 Neubert 来说,持续集成可确保该通道的行为保持正确和一致,因为他的团队正在改进其代码,从而提供一种“极具价值”的保障。“真正的危险是错过了一些东西,或者犯了一个小错误。”他说。

例外

目前已经存在各种各样的持续集成服务。这些产品包括开源无人机,以及诸如 CircleCI、Codeship、GitLab、Shippable 和 Travis CI 等商业



图片来源:the Project Twins

选择,它们都提供基于所需测试行为、用户数量以及项目是公是私的定价等级。例如,Travis CI 对开源项目是免费的,但私人项目的费用则是每月 69 美元。Shippable 为公共项目提供免费的基本服务,但每月收取 25-150 美元的费用支持私人项目和更大的计算能力等性能。

然而,研究人员应该考虑什么是合适且值得投资的。并非每个项目都需要持续集成,并设置和配置一种具有挑战性的服务。费城宾夕法尼亚大学数据科学博士后 Daniel Himmelstein 说,如果服务需要与使用权限存在法律限制的软件或数据进行交互,那么就会出现更多的困难。而且,代码通常仅用一次,这样降低了成本。“对于日常的研究编码,代码数量仍未大到可以使持续集成有价值。”加州大学圣地亚哥分校的高性能计算专家 Andrea Zonca 说。他在发布代码时使用 Travis CI,但他所写的大部分代码都是为了自己的一次性使用,且不会再次执行。

如果代码不断更新则需要重复测试,那么计算成本也会增加,这就是为什么 Neubert 的实验室在代码修改后只测试最关键的数据分

析的原因。

尽管面临这些挑战,但持续集成服务往往会提高代码质量,德国弗莱堡大学的生物信息专家 Bjorn Grüning 说,尤其是在类似 Galaxy 的大型项目上,这是 Grüning 和其他 160 多人一起研发的一个生物信息学工具包。

据 Grüning 介绍,持续集成缩短了批准 Galaxy 项目的周转时间,在提交新功能和补丁时给了程序员更大的信心。他说,在没有这些服务之前,类似项目中的研究人员在验证合作者提出的每个新特征时经常会脱离实际,因为他们没有时间。

一些研究人员使用持续集成实现非编程任务的自动化。今年 4 月,作为一个研究生态系统如何随时间变化的项目的一部分,佛罗里达大学生态学家 Ethan White 帮助配置了 Travis CI 以自动更新新的野外或气象站数据的图表,从而让该研究团队每个月可以节省 5 个小时。

持续集成帮助 Himmelstein 自动修改科学论文引用和网页文本或代码更新。他说,如果没有持续集成,维护人员可能会“变得懒惰,不能每次都更新手稿”。

初始化

无论由第三方托管还是在用户自己的机器上,持续集成服务都是由一组自定义的指令控制的。这个配置文件定义了要运行的任务,并根据需要运行它们的正确环境(操作系统和软件库)设置服务端。然后,该服务在设置时间接收到一个代码或数据更新时便会执行这些指令。宾夕法尼亚大学生物信息学家 Casey Greene 利用持续集成重新进行数据分析,他已经验证了当今许多最受欢迎的服务。“关于这些服务的好消息是——它们非常相似。”他说。

但也存在着一些细微的差别,例如,目前用户可以运行的并行工作的数量,或是可以运行它们的计算能力。“我鼓励人们挖掘每项服务的极限,以确保它们与工作流兼容。”Greene 建议。

Greene 说,尽管目前科学领域所采用的持续整合规模仍然很小,但它在增长,更多研究人员都应该加入进来。不过,他坦言提升速度需要时间,但通常付出的努力是值得的。“分析数据的科学家应该把它放在他们的工具箱里。”(晋楠编译)

吃药能治昏睡病

研究者希望管理机构能批准这种快速而简单的疗法



卫生工作者在刚果民主共和国的一个偏远村庄里对人们进行昏睡病筛查。 图片来源:Neil Brandvold

锥虫能感染人类中枢神经系统,患者通常会出现意识错乱、日间嗜睡、失眠症及多种精神症状,包括躁狂性发作和攻击性等,如果患者没有得到及时治疗就会陷入昏迷,甚至死亡。多年来,治疗昏睡病的唯一疗法就是以毒性砒霜为基础开发的药物,但这种药物常常会使 1/20 的患者死亡。

2009 年,研究人员开发出一种更安全的疗法选择,即硝咪莫司-依氟鸟氨酸组合性疗法(NECT 疗法),这种疗法包括药物服用和 14 针的静脉内药物注射。在采用这种疗法后,研究人员 50 年来首次使昏睡病每年新增病例降低到 1 万例以下。

据世界卫生组织数据,目前每年昏睡病的

发病数量维持在 2200 例左右,但在治疗过程中,患者需要输液,同时还需要接受脊椎抽液以评估疗法的效果,对于很多无菌设备、电力和医生短缺的地区仍然存在一定障碍。

开发 NECT 疗法的机构是瑞士日内瓦非营利研究组织——被忽略疾病药物创新计划(DNDi),它们随后开始研究开发另一种更好的疗法。2007 年,研究人员发现了化合物非昔硝唑——赛诺菲公司开发的一种药物,在与赛诺菲签订相关协议后,DNDi 在相关的临床试验中使用了该药物。据估计,这种疗法从开发到获批将花费约 5000 万美元,只是制药公司开发新药成本的一小部分。

赛诺菲公司不久将向欧洲药品局提交药物审批申请,该局的审批将会为这一疗法在刚果民主共和国的使用铺平道路。DNDi 的医学总监 Nathalie Strub Wourgrat 表示,预计这种药物在明年年底前将会获得绿灯,由于其是一种口服药物,患者可以在家进行治疗,将大大节省患者治疗的时间以及开支。

不过,Büscher 认为,在家中服药治疗或许具有一定的危险性,因为对于药物非昔硝唑没有反应的人群,如果没有得到医务人员的及时救治很可能会死于昏睡病,因此患者应在卫生工作者的指导下服用药物。而且,他还建议向患者提供返回医院进行治疗的动力,比如提供金钱补助或食物补助等。“这是一次成功,但不是终结。”他说。

这种口服非昔硝唑疗法或许仅仅是开始,目前 DNDi 研究人员正在进行大量研究,希望能开发出新型疗法让患者仅进行单次剂量的药物服用就能够彻底治疗昏睡病,同时他们也希望开发出比非昔硝唑疗效更可靠的药物疗法。(张章编译)

科学线人

全球科技政策新闻与解析

欧盟出台“地平线 2020”新一轮支出方案



近日,欧洲委员会公布了欧盟科研创新资助计划“地平线 2020”2018-2020 年度支出方案,总金额 300 亿欧元,其中包括 27 亿欧元用以启动欧洲创新理事会。

据悉,“地平线 2020”是欧盟科研与创新资助计划,预算将投入 770 亿欧元,支持欧洲卓越科学行动,并促成了发现系外行星和引力波在内的突破性成果。未来 3 年,欧盟的科研资助将聚焦于能带来更大影响、关注度更少但更关键的课题,例如移民、安全、气候、清洁能源及数字经济。

例如,低碳和气候相关研究将获 33 亿欧元拨款;投入 10 亿欧元发展循环经济;欧洲工业和服务业数字化及转型将有 17 亿欧元;22 亿欧元拨款将投入清洁能源 4 个相关领域的项目:可再生能源、能效建筑、电动运输和储能方案,其中,2 亿欧元支持研发生产欧洲下一代电池。

“地平线 2020”还将进一步推动突破性、市场驱动性创新。欧盟委员会研究、科学与创新委员卡洛斯·莫达斯说:“人工智能、基因、区块链等是当今最具前途的创新核心。欧盟委员会正努力为欧洲众多创新者提供一个跳板,助力他们发展成为世界领先者。”

目前,欧盟还设立了一系列经费奖励方案,以促进其科研研发活力。例如,“地平线 2020”还将通过“突破挑战”奖项,鼓励科学家关注民众面对的迫切问题,研发突破性技术方案。该项目还将继续资助“好奇心驱动的科学研究”(“蓝天科学”)。欧盟将为卓越的科研学者提供近 18.6 亿欧元的经费支持。玛丽居里行动计划为各个职业发展阶段的研究人员提供奖学金资助,未来 3 年总预算将达 29 亿欧元。

此外,2018 年至 2020 年,“地平线 2020”将拨款 4.6 亿欧元特别支持尚未充分参与欧盟成员国内框架计划协约国;并拨款 20 亿欧元支持科学开放,其中 6 亿欧元资助欧洲开放科学云、数据基础设施以及高性能计算机。(唐凤)

专家建议美国大幅调整动物研究规则



图片来源:ROBERT F. BUKATY/AP IMAGES

数年来,研究者和大学一直在抱怨美国研究用动物福利管理条例的混乱。例如,涉及兔子和大型哺乳动物的研究主要由美国农业部(USDA)进行监管,由联邦政府资助的啮齿类动物和鸟类研究,则由美国国立卫生研究院(NIH)用不同的条例管理。

与此同时,许多私人资助的动物研究受到的监督相对较少。“这就像是一床用碎布拼成的床。”美国医学院协会(AAMC)首席科学官 Ross McKinney 说。

现在,AAMC 和其他 3 个团体正在推动该国对动物研究规则的全面改变。在近日发布的一份报告中,专家呼吁将所有监督都转移到一个机构,减少实验室检查的频率,并在制定新规则方面给予研究人员更大的发言权。McKinney 指出,这些变化将确保“我们在保护研究动物”,“但我们希望这样做的方式是一致、连贯和有效的”。

USDA 数据显示,2016 年超过 80 万动物被用于研究,包括猫、兔子和灵长类动物。(这还不算数百万老鼠和鸟类)参与该报告研的华盛顿大学实验动物监督负责人 Sally Thompson-Iritani 表示,在现行制度下,“我们花了大量时间用于记录档案”。

Thompson-Iritani 举例称,一个大学委员会必须每 6 个月检查一次动物设施,这一过程需要 4 个月,USDA 要求大学每年审查协议,而 NIH 则要求每 3 年进行一次。

于是,该报告呼吁放宽许多要求,建议每年只要求大学进行一次设施检查,每 3 年进行一次协议审查。它还建议国会修改《动物福利法案》,废除政府官员每年进行的现场检查,转而集中检查有问题记录的设施。而为了减少文书工作,报告敦促大学委员会对某些类型的实验进行豁免。

但也有人不担忱,这样的改革将让研究人员默认采取最低严格限度的措施。(张章)