

深入浅出

AIOps



卷首语

智能时代的新运维

InfoQ 主编 徐川

今年是 CNUTCon 的第三年，极客邦科技 InfoQ 中国的第十年。有一些 CNUTCon 的老用户可能会看到，今年 CNUTCon 的主题和去年前年不一样，之前都是容器，而今年我们定位是智能时代的新运维。

要说原因也很简单，作为一家技术媒体，我们的使命是推动软件技术的发展，去年、前年我们看到了容器为运维领域带来的种种变革和创新，而今年，到这个节点我们看到 AI 可以很好地解决运维中的一些问题。

我们希望不管是极客邦科技提供的课程、视频、新闻、文章、直播还是会议，都能向整个的开发者社区提供最为前沿的理念和实践。

我们自己定位是一家内容驱动的公司，那我们怎么做内容了？这里我简单和各位分享下我们重点参考的技术采用生命周期模型。

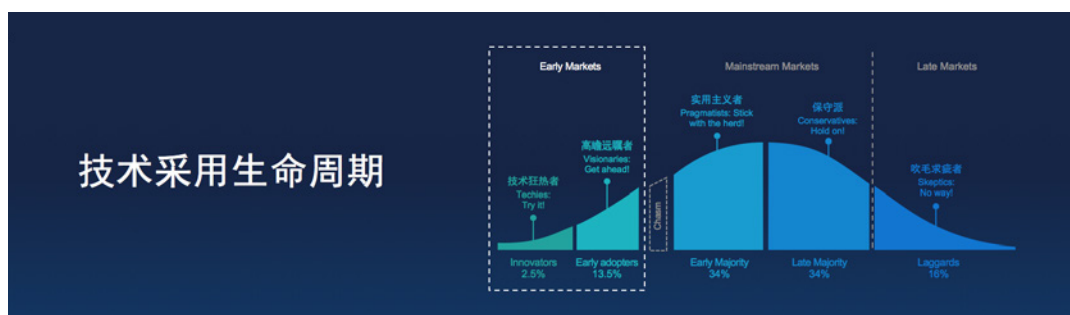
技术采用生命周期

「用来衡量用户对某项新技术接受程度的模型」

简单来说，技术采用生命周期是一个用来衡量用户对某项新技术接受程度的模型，这个理论最早源于 1943 年对玉米新品种的扩散行为研究，而后 1962 年因为《创新的扩散》一书的出版逐渐变得流行起来。

这一曲线将用户采用新技术的过程分为五个阶段，分别包括创新者、早期采用者、早期大众、晚期大众与落后者。简单来说，一个技术从开始到最后流行都会逐步经历这五类人群，比如一开始的创新者，在新技术来的时候，他们只要觉得东西不错，就会毫不犹豫的采用。但越往后走，大家越保守。

任何一个产品从面世，到最后发展成熟，都会经历这些人群，以及这些发展阶段。比如 OpenStack 的发展。



图中大家可以看到，在早期采用者和早期大众之间，有一个巨大的鸿沟，很多技术很多产品也都是因为没有走出这一个跨过这一鸿沟而消亡，也有的人叫它死亡之井。所以 InfoQ 重点关注的是前两个阶段的技术，我们通过会议、线上内容、课程等来报道这两个阶段的技术，希望能够被更多的用户知晓并采用，以帮助这些技术跨越鸿沟。

所以简单来说，我们做内容时参考的模型就是这个技术采用生命周期

模型，我们关注处于创新和早期采用阶段的新技术，以帮助他们跨过死亡之井（Chasm）。

在这里，我也总结了从我们看来的运维领域重点的几个技术。他们都没有跨越鸿沟，我们也是在重点关注这些技术点。



接下来，我结合今天的主题，和大家重点聊聊 AIOps。前面和大家介绍了这么多，就是想告诉大家我们做内容的思路，以及为什么我们要定智能时代的新运维这个主题，是因为我们的编辑从中看到了一些发展趋势。

前两天，我在微博上和一个粉丝还讨论，它说为啥你们又搞出一个新名词，上一个新名词 DevOps 我还没消化明白了。我说，这不是新名词，我们相信 AIOps 是运维发展的必然趋势。如果我们不抬头看路，瞄准未来，那可能就是井底之蛙了。

为什么这么说了？我们先来回顾下整个运维的一个发展历史，运维人员一开始的称谓其实是系统管理员，然后逐步过渡到后来的运维，再后来又有了 DevOps 或者 SRE 的称谓。我们其实可以看出来，这些年，随着互联网技术的逐渐成熟，企业对于运维的能力要求越来越高，整体来说，就是要让运维具备开发的能力。



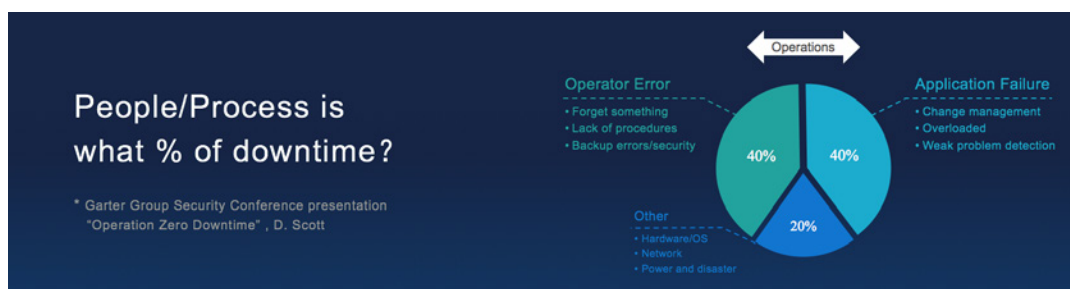
为什么会有这样的趋势呢？我认为是整个互联网应用的运维越来越复杂了，动不动就是几万个应用，几个 9 的可靠性，运维管理的复杂度和难度成倍数的增大，业务的告诉发展倒逼运维的迭代。

这些年，运维平台的演进我觉得大体是这样的：



这两年炒的比较多的概念还有一个就是自动化运维。借着这个大会，我采访了很多自动化运维的团队，也都问了大家说这么多年自动化运维实施下来，都解决了他们哪些问题。得到的答案出奇的一致，基本围绕在效率这个词上。

记得在 2001 年的时候，Gartner Group 有一个调查显示在 IT 项目经常出现的问题中，源自技术或产品（包括硬件、软件、网络、电力失常及天灾等）的问题只占 20%，但流程失误方面却占 40%，人员疏失方面也占到了 40%。



这些年来，企业通过自动化运维平台以及 DevOps 等协作理念其实是逐步在解决了 Gartner 提到的流程失误和人员疏忽相关的 80% 的问题。虽然目前没有具体的统计数据，但可以确认的是，这一问题得到了有效解决。

但另外一个值得注意的点是当前的 IT 项目基础设施环境与 5 年前已经没有办法同日而语，更不用说 10 年前。近几年，随着云计算、微服务等技术的流行，以及互联网业务的迅速发展，运维人员要关注的服务数量也呈现了指数级增长，自动化运维虽然提升了效率，解决了一部分问题，但也遇到了新的难题，比如面对繁多的报警信息，运维人员应该如何处理？故障发生时，又如何能够迅速定位问题？

「AI + Ops」？

Gartner 在 2016 年提出了「AIOps」概念，并预测到 2020 年 AIOps 的采用率将会达到 50%

这个时候，恰好历史进入了人工智能时代，那上面的这些问题可否通过『AI + Ops』的这种跨界创新的方式来解决呢？于是 Gartner 在 2016 年时便提出了 AIOps 的概念，并预测到 2020 年，AIOps 的采用率将会达到 50%。简单来说，AIOps 就是希望基于已有的运维数据（日志、监控信息、应用信息等）并通过机器学习的方式来进一步解决自动化运维没办法解决的问题。

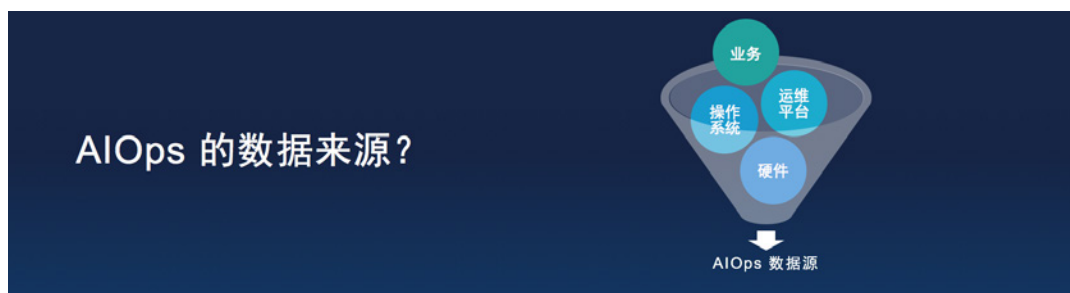
如果说要下一个定义的话，我认为是这样：

「AIOps」

通过人工智能的方式，进一步提高运维效率，包括运维决策、故障预测、问题分析等

那紧接着又有一个问题，人工智能需要有数据，AIOps 的数据从哪里来？我想各位也一定能列举出很多，比如各种各样的监控数据、各种各样

的日志数据，各种各样的硬件数据。总结起来，有这几类：



当然，这不是我凭空说的，这是百度同学总结的，晚上我们有 AIOps 的圆桌讨论，到时候可以过来聊聊。

你肯定会问，说 AIOps 有怎么样的应用场景呢？这里我来说几个例子，阿里巴巴很早之前就在利用机器学习的方式来预测磁盘故障，而百度的智能运维，这个一会百度 AIOps 的同学也会分享，应用也非常广泛，包括性能优化、故障分析、流量调度等等。



那现在有哪些 AIOps 的落地实践呢？这里我分享一下我看到已经开始实践的企业，并且在网上能查到相关资料的。供各位参考。



目录

09 AIOps 是什么？它与 AI 有什么关系？

12 赵成：回顾运维的发展历史，我相信 AIOps 是必然趋势

20 到底应该如何理解 AIOps？又如何落地 AIOps？

27 虚拟座谈会：聊聊 AIOps 的终极价值

AIOps 是什么？它与 AI 有什么关系？

作者 薛命灯



现如今，AI 这个词已经被玩坏了。很多公司都声称自己在做 AI，但其实并没有。不过有另外一种新兴的 AI，各种类型的 IT 企业倒是可以尝试，而且完全不需要人工参与。

AIOps，也就是基于算法的 IT 运维 (Algorithmic IT Operations)，是由 Gartner 定义的新类别，源自业界之前所说的 ITOA (IT Operations and Analytics)。我们已经到达了这样的一个时代，数据科学和算法正在被用于自动化传统的 IT 运维任务和流程。算法被集成到工具里，帮助企业进一步简化运维工作，把人类从耗时又容易出错的流程中解放出来。人们不再需要在遗留的管理系统中定义和管理无穷无尽的规则

和过滤器。

在过去的几年间，一些新技术不断涌现，利用数据科学和机器学习来推进日益复杂的企业数字化进程，“AIOps”（Algorithmic IT Operations）因此应运而生。Gartner 的报告宣称，到 2020 年，将近 50% 的企业将会在他们的业务和 IT 运维方面采用 AIOps，远远高于今天的 10%。

为了更好地理解 AIOps 和 AI 的区别，我们需要从头说起。

AI 简史

AI 一词用于描述机器（或软件）模拟人类认知的过程。也就是说，机器学习像人类一样思考。40 年代，Alan Turing 掀起了 AI 热潮，但受限于计算机的计算能力，也只发展到今天的这个阶段。

问题是，我们为什么要让机器模仿人类？而为什么有些 AI 应用程序会比其他的更成功？发展 AI 的目的在于解决人类的问题，所以我们会看到像自动驾驶汽车、行为分析这类复杂的解决方案。

话说回来，IT 运维环境有一些不一样的地方。我们不会直接管理人类，我们与应用程序和基础设施打交道。而且它们可能更加复杂和不可预测，因为它们不是人类。

人类思维与机器思维

AIOps 的不同之处在这里体现出来。AIOps 的解决方案专注于解决问题，而且是通过使用基于算法的技术来高度模仿人类（而且以更快的速度和更大的规模）。算法的效率提升了 AIOps 的价值，而相对于人类的智慧——虽然是无限的，但不如机器来得高效。

当然，人类也能进行高效的 IT 运维。AIOps 的目的是为了让我们的生活变得更美好，但是当人类与 AIOps 参合在一起，它们之间的界限就会变得模糊。高级的 AIOps 会使用神经网络技术，它会向运维人员学习，然后尝试消除无聊的重复性劳动。

未来的公司

为什么公司需要 AIOps？现代的 IT 环境已经无比的复杂，而且千变万化，需要我们花费大量的时间和资源去监控、去诊断问题、去解决问题。很多公司处于被动的地位。但是如果他们使用了 AIOps，他们就可以利用先进的算法，花更多时间在其他更有意义的工作上，而不是重复地解决相同的问题，或者花时间管理规则和过滤器。

我们所说的规则，可以把它们简单地描述为“如果是这样那么就这么做”，它们能够应付简单的场景，但是很难扩展。相反，算法和机器学习提供了更加灵活的表达方式，不仅强大，而且健壮，能够应付不断变化的需求。这将带来更高的效率和更低成本。对于厂商来说，他们面临的挑战在于将整个技术方案打包，避免把用户暴露于底层的复杂性当中。光是提供工具是不够的，企业需要招聘数据科学家而不仅仅是工程师。

前行之路

借助智能算法的技术优势，原先人工需要几个小时完成的任务现在通过自动化可以在几秒钟内完成，而且能够得到更好的结果。传统的 IT 运维需要管理大量的告警，极大地分散了企业的注意力，他们需要花很多时间解决无聊的问题，没有时间用于创新。使用 AIOps 可以解决这些问题，把运维人员从纷繁复杂的告警和噪音中解脱出来。各个行业的企业正在采用 AIOps，他们使用这项技术来改进客户的数字体验——银行、娱乐、交通、零售，甚至政府。

尽管 AIOps 还是一个新名词，但并不代表它只是未来的一种趋势而已。在这个数字的年代，任何使用传统技术来管理机器数据的组织要么忽略了信息的价值，要么已经让他们的运维团队不堪重负。随着数据的暴涨，CIO 们应该快速拥抱 AIOps。传统 AI 仍然会在某些领域发挥它的作用，而 AIOps 将为企业带来最直接最深远的价值。

赵成：回顾运维的发展历史，我相信 AIOps 是必然趋势

作者 郭蕾



毫无疑问，运维技术的发展已经进入了深水期。随着 Docker、OpenStack、Puppet 等技术的流行，以及 CI/CD、DevOps 等理念的落地生根，自动化运维的发展迎来了小高潮。整体来看，自动化运维平台帮助提升了运维的效率，并减少了因人工和流程操作而引起的运维故障。

记得在 2001 年的时候，Gartner Group 有一个调查显示在 IT 项目经常出现的问题中，源自技术或产品（包括硬件、软件、网络、电力失常及天灾等）的问题只占 20%，但流程失误方面却占 40%，人员疏失方面也占到了 40%。这些年来，企业通过自动化运维平台以及 DevOps 等协作理念逐步解决了 Gartner 提到的流程失误和人员疏忽相关的 80% 的问题，虽然

目前没有具体的统计数据，但可以确认的是，这一问题得到了有效解决。

但另外一个值得注意的点是当前的 IT 项目基础设施环境与 5 年前已经没有办法同日而语，更不用说 10 年前。近几年，随着云计算、微服务等技术的流行，以及互联网业务的迅速发展，运维人员要关注的服务数量也呈现了指数级增长，自动化运维虽然提升了效率，解决了一部分问题，但也遇到了新的难题，比如面对繁多的报警信息，运维人员应该如何处理？故障发生时，又如何能够迅速定位问题？

当企业遇到这些新的问题却无从下手时，恰好历史进入了人工智能时代，那上面的这些问题可否通过『AI + Ops』的这种跨界创新的方式来解决呢？于是Gartner在2016年时便提出了AIOps的概念，并预测到2020年，AIOps的采用率将会达到50%。简单来说，AIOps就是希望基于已有的运维数据（日志、监控信息、应用信息等）并通过机器学习的方式来进一步解决自动化运维没办法解决的问题。

就目前来看，国内的百度、搜狗、宜信、阿里巴巴都已经探索尝试了AIOps，并且取得了不错的收益。在2017年InfoQ举办的CNUTCon全球运维技术大会上，也有不少AIOps相关的议题，甚至会议主题也从去年的容器生态迭代到今年的智能时代的新运维，感兴趣的读者可以关注。

那对于AIOps这个新名词，它又会涉及到哪些新技术？从运维的发展角度看，为什么说AIOps是必然趋势？它与自动化运维之间会有什么样的关系？InfoQ记者就这些问题采访了美丽联合集团运维经理赵成。

InfoQ：如何理解AIOps？AIOps会涉及哪些技术？这又是一个新名词吗？

赵成：我觉得理解AIOps之前，还是先理解下AI、机器学习、深度学习这样几个概念。如果用一张图，来表示，就是图1。

同时，这张图的[来源文章](#)，也建议看一下，讲的还是比较清晰的。

简单来说，AI，人工智能是一个广义概念，最早期提出来的时候，人们的愿景是希望AI能够完全具备人类智慧，这属于“强人工智能（General AI）”。但是除了在科幻电影和科幻小说外，在现实中还没有实现，可能

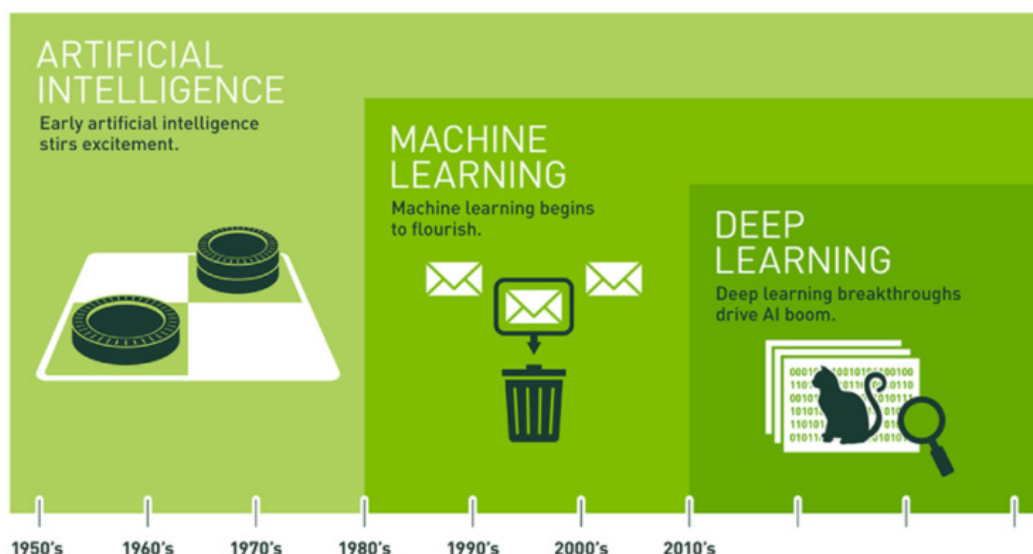


图 1

不仅仅是要求对计算机技术要求极高,对生物医学技术也会有很高的要求,因为前提是我们得看懂人类大脑运转的每一个细节吧。

但是过程中,研究人员发现我们虽然没法让机器完全具备人类一样的智慧,但是在非常具体和特定的领域,机器是可以做的比人类更好的,比如图形图像处理、语音识别等等,这些人工智能的应用,称之为“弱人工智能(Narrow AI)”,这些应用的实现手段,就得益于机器学习算法长足的进步,而深度学习又是机器学习领域很精深的一部分。

所以,我们现在提到的AI,更多的是依赖机器学习(包含深度学习)算法的实现的AI场景,或者说机器学习算法只是实现AI的其中一种手段。

了解了上面的概念,再回到AIOps上来,拆分为AI + Ops会准确一些,也就是Ops与AI相结合可以做的事情。Gartner的定义是Algorithmic IT,而不是Artificial Intelligence,我起初觉得也不是很合理,但是我认真理清整个关系之后,我觉得这样定义也没有问题。

最后,我觉得定义如何到没有必要纠结,因为不管AIOps里这个AI到底是Algorithmic IT还是Artificial Intelligence,最终,我们根本上使用的,还是机器学习算法这个手段。

AIOps涉及的技术,从AI的角度,主要还是机器学习算法,以及大

数据相关的技术，因为涉及到大量数据的训练和计算，从 Ops 的角度，主要还是运维自动化相关的技术。另外 AIOps 一定是建立在高度完善的运维自动化基础之上的，只有 AI 没有 Ops，是谈不上 AIOps。

InfoQ：你认为 AIOps 是运维发展的必然趋势吗？从手工运维，到自动化运维，再到现在的 AIOps，谈谈你理解的运维发展趋势？

赵成：必然趋势。一个很明显的规律，凡是能让我们的生活变得更美好、更简单、更方便的技术，一定会具有强大的生命力，也必然会成为发展趋势，而 AI 正是这样的技术之一，AIOps 又是其中的一个专业领域。

运维的发展变化，我的感受，是随着业务和技术发展变化的，根本上还是业务驱动和倒逼出来的。

比如 2008 年 -2010 年，我接触的是电信级软件的开发和维护，那个时期的软件有这么几个特点：

业务场景和形态上，相对固定，变化不大；

软件是分层架构，模块数量固定，架构上基本不会有太大的变化；

研发流程和规范非常严格，基本交付上线之后不会有太大的功能和性能问题，线上出问题会面临非常严格的处罚；

软件交付周期长，大的变更半年一次，小变更至少 3 个月，这期间研发会将发布脚本以及产品升级文档制定的非常详细，我们称之为 Step by Step，而且也是经过严格验证和测试的；

研发维优团队支持，一旦线上出现软件方面的问题，一线维护直接将问题转回研发进行定位处理，而且有严格的 SLA 约束。

这种情况下，变化不大、研发可以搞定绝大部分软件层面的事情，所以我们会看到这个时期的运维更多的是网络、硬件、系统方面的维护职责，这个是由那个时期的业务特点，以及软件架构特点所决定的。

当然这并不代表那个时候业务的技术难度和复杂度不高，反而技术门槛是相当高的，协议之复杂，网元交互之多，架构图和交互逻辑画出来也是相当复杂，那个时候我们答辩和述职必做的事情就是看网讲网，要对所有的周边网元以及交互关系讲清楚。

之后，随着互联网业务的高速发展，业务场景上丰富多彩，复杂和多变，新业务和新场景也在不断涌现，为了快速验证产品和需求的方向是否有效，就要快速试错，对迭代开发和交付效率有了极高的要求，所以技术界逐渐催生出了服务化这样的软件架构，以及持续交付过程，同时随着业务体量快速膨胀，衍生出对稳定性有极高的要求，这时我们现在长听到的全链路跟踪、容量评估、限流降级、强弱依赖等稳定性的解决方案就涌现出来。

这时对于开发的诉求是能够将更多的精力放到需求实现上，而因服务化带来的大量的应用管理、持续交付、监控、稳定性、成本控制等非功能性体系的建设和保障就需要有专门的团队来做，这时对于运维的诉求也在悄然发生着变化，所以这也是为什么这个阶段会涌现出 DevOps、技术运营、PE、SRE 等等对运维重新定位的词汇。当然这个过程中，因为云计算发展，传统的网络、硬件和系统维护的职责在逐渐的被弱化，也在逼迫着运维的关注点从底层转向应用和业务层面。所以，我们看到就在近 2-3 年，自动化、发布系统、稳定性平台这些系统成为了运维团队重点关注和建设的部分。

这里就有一个趋势，就是 SpringCloud 和 Cloud Native，SpringCloud 让当前服务化的开发变得越来越方便和高效，而 Cloud Native 在打造应用的基础设施方面也已经取得了长足的进步，而且已经有了 CNCF 这样的组织在驱动响应的标准和体系建设。设想一下，后面从业界的角度，如果 Spring Cloud 成为微服务的开发标准，Cloud Native 成为应用的运维标准，是不是又会驱动着一波运维的转型和升级呢？

回到 AIOps 上来，当前这个阶段，现实情况，系统里面已经有大量软硬件模块、日志、监报告警指标也纷繁复杂，一方面是无法在问题萌芽状态就发现问题，无法提前做出预判，另一方面是发生了问题又无法快速确定根因，造成持续的资损。技术发展上，随着计算能力、数据量的积累、以及机器算法的进步，如何更加高效的开展 Ops 这个问题就摆在我们面前，AIOps 的模式应运而生。

所以，运维一步步发展到当前这个状态，根本上讲还是业务高速发展

倒逼出来的，同时，从手动运维到运维自动化，再到 AIOps，这个过程根本上是在朝着如何更加高效运维的趋势在发展。

InfoQ: AIOps 的出现是为了解决哪些问题？这些问题运维自动化无法解决吗？

赵成：主要还是解决复杂环境下问题的快速发现甚至提前预判，以及出现问题后的如何在复杂的告警、报错和日志中快速进行根因分析。

运维自动化无法解决吗？我的理解，AI 和 Ops 要解决的还是两个层面的问题，可以类比到人，AI 相当于人的大脑，我们手脚和躯干是执行系统，大脑负责决策判断，手脚躯干负责完成大脑下发的动作指令。对应到运维上面，AI 要解决的是怎么快速发现问题和判断根因，而问题一旦找到，就需要靠我们高度完善的自动化体系去执行对应的运维操作，比如容量不够就扩容、流量过大就应该触发限流和降级，等等。

AI 是能够让 Ops 执行的更加高效的强大助推力，下面是我之前整理出来的，我理解的 AIOps 的体系和建设思路（见图 2）。

InfoQ: 落地 AIOps 的前提条件是什么？什么样的团队适合落地 AIOps？

赵成：AIOps 的首要前提条件，一定是先要有高度完善的运维自动化，自动化都没做好前，先不要玩 AI，千万不要本末倒置。

从 AI 的角度，应该有三个方面的充要条件：机器学习算法、计算能力（类似于 GPU 的高性能设备）、海量数据。

先看算法，这个基础，在 AI 中我们使用到的各类算法，比如基于指数平滑的二次平滑、三次平滑算法，基于分解的傅里叶分解、小波分解算法等，基于深度学习的前馈神经网络、循环神经网络 RNN 算法等，这些算法早就已比较成熟了，并大量使用在其它的研究领域，比如我们熟悉的图形图像处理、语音识别，还有在医疗、电力以及通信行业的应用。所以在算法上，我们很早之前应该具备了这方面的理论基础。

但是 AI 为什么这几年突然火起来，或者在应用上有了长足的进步，很大原因就是计算能力提升，海量数据积累起来了。比如随着硬件计算

能力的提升，有了 GPU 这样的超高性能计算设备，同时还有云计算这样规模的基础设施支持；再就是，最近这些年随着互联网业务的高速发展，各行各业都积累了海量的现实的数据。

比如吴恩达教授非常著名的识别猫的实验，一方面吴教授有自己非常牛的深度学习算法，另外一方面，基于 16000 个处理器的计算能力搭建起了深度学习的平台，然后基于 Youtube 中千万级别的猫的视频图像进行算法的学习和训练，最终可以非常精确的从图像中识别出猫。

回到 AIOps 上面来，看这个三个条件：

算法还是那些算法，不过得要有相应专业能力的团队，如果是纯应用，我觉得运维团队倒是可以自学一下，但是不管怎么样，这个还是有一定门槛，需要大量的学习和能力提升。

数据就是要靠线上运行的真实数据和日志，所以必须要有大量的数据积累

计算能力上，目前看到我们基于大数据技术的数据处理能力已经足够，因为毕竟不是像图形图像这样的复杂计算场景。

从上面三个条件看，也就不难理解，AIOps 做的比较超前的为什么都是那些国内外的大厂，因为有能力、有足够的资源、有足够的技术实力、有足够的数据，最关键的是足够复杂和变态的业务场景以及运维场景，在倒逼着 Ops 往这个方向上走。

至于什么样的团队适合落地，我暂时没想到什么标准，不过还是那个建议是，先尽快做好自动化，把基础打好，AI 的学习上做一些储备，当业务复杂度和体量到达一个

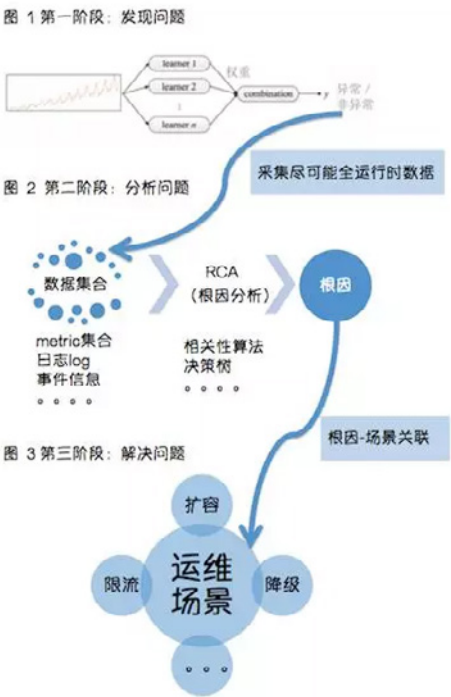


图 2

量后，会自然倒逼着运维往这个方向发展，千万别自动化还没做完善，就跟风搞 AIOps。

InfoQ: AIOps 中的数据是怎么来的？数据是必要的吗？

赵成: AIOps 中的数据必须是线上产生的现实场景下的运行数据，不管是底层硬件和系统层面，还是应用和业务层面，以及运维的操作记录日志，要尽可能全面的数据。这些数据一方面要做算法模型的训练，让算法能准确识别问题，一方面还要在问题分析时做根因分析使用。

数据是必要的，准确说是必需的。目前 AIOps 中，就异常发现来说，针对不同的应用场景，应该使用哪种算法模型，这个还是有一定挑战的，所以起初可能会同时使用多种算法同时运行，这时就需要大量真实的数据去验证算法运行的情况，同时做一些参数校正，也就是我们所说的训练的过程，最终我们根据跑出来的结果准确度选择合适的算法，或者设定相应的权重。所以，机器学习算法是否有效是离不开大量的真实数据的训练的。

InfoQ: 可否谈谈你们的 AIOps 落地场景？

赵成: 这块我们还在实践中，一块是异常检测，做一些关键监控 Metrics 的曲线监控，这块用到的基本是常见的指数平滑、3-Sigma 算法等。另一块是根因分析，在服务化的架构中，最头痛的还是出现了故障，无法快速的定位原因。大致思路是，根据全链路跟踪系统的每一次请求的依赖关系，做调用的关联度分析，当一个模块出现问题时，会同时导致依赖这个模块的所有模块都会告警，甚至还有业务层面的告警，这时就需要快速的根因分析，确定问题在哪儿。

嘉宾简介

赵成(谦益)，美丽联合集团运维经理，负责美丽联合集团（原蘑菇街、美丽说）运维团队管理及运维体系建设工作。拥有近 10 年研发和运维经验，见证和参与了多个电信级和互联网产品从从无到有的建设，从微量到海量的成长过程，拥有非常丰富的电信级和互联网业务研发和运维经验。目前专注于运维创造价值，以及云计算和 AI 时代运维的转型和突破。

到底应该如何理解 AIOps？又如何落地 AIOps？

作者 郭蕾



近年来，人工智能技术备受关注，将 AI 引入 IT 运维领域，AIOps 的概念由此应运而生。Gartner 的报告宣称，到 2020 年，将近 50% 的企业将会在他们的业务和 IT 运维方面采用 AIOps，远远高于今天的 10%。尽管 AIOps 还是一个新名词，但 并不代表它只是未来的一种趋势而已。在这个数字的年代，任何使用传统技术来管理机器数据的组织要么忽略了信息的价值，要么已经让他们的运维团队不堪重负。

那就当下而言，我们应该如何理解 AIOps？AIOps 应该如何落地？能否通过 AIOps 支持更好的运营？带着这些问题，我们采访了宜信技术研发中心高级架构师张真，请他从宜信近几个月落地 AIOps 的角度聊聊

他的想法和洞见。另外，张真也会在 CNUTCon 全球运维技术大会 上分享他们的实践案例。

InfoQ: 你是如何理解 AIOps 的? AIOps 的关键点是什么?

张真: 我认为 AI 的生态体系与大数据类似，存在两种基本角色：AI 科学家和 AI 领域工程师（FE）。前者推动 AI 科学的发展，创造新的 AI 知识体系；后者是将 AI 知识运用到生产生活的某个领域，创造现实价值。AIOps 正是将 AI 技术应用到 IT 运维领域，帮助变革运维模式，提升效率和创造现实价值的“工程化”过程。

从实施角度，AIOps 这个词本身就体现了两个关键点。其一，Ops 代表运维的场景，这是主旨，识别什么样的场景存在哪些痛点，AI 可以帮助解决；同时也要清楚认识目前的 AI 技术擅长什么，不擅长什么，有哪些限制，切忌凡事尽 AI。其二，AI 作为前缀代表技术，这是手段，AI 技术门类很多，选择合适的，正确的技术去解决真正的问题，是需要切实履行的原则。此外，要从实际出发，考虑投入与产出。

从技术角度，也有两个关键点。首先 AI 的目标是系统类人化，而 AIOps 是将运维系统类人化。它的技术栈应该涵盖三个基本特征：类人交互，主动决策，理解执行。这是与自动化的本质区别。其二，与 DevOps 工具链深度集成是必由之路。我认为 AIOps 不是要替代现有的工具链，而是通过类人化提升“智慧”，实现 SRE 甚至超 SRE 的效果。要达成这个目标，AIOps 就要“学习，了解”这些工具，并且更好的“使用”这些工具，这个过程就是深度集成，它的核心是对这些工具 API 的自主认知和自主使用。

InfoQ: AIOps 和 DevOps 有什么关系吗? 可否聊聊你看到的运维理念的演进过程?

张真: 正如刚才提到的，AIOps 在技术层面要对 DevOps 工具链进行深度集成。另外，我也想从发展历程角度谈谈它们的关系，其实我也是伴随着这些发展阶段逐步成长起来的。这应该也是个人运维理念的演进过程吧。个人认为 运维经历了 4 个基本发展阶段：

早期工具时代：这个时期是 IT 运维的软件工具，流程初始化的时期，工具的目标仅仅只是计算机化，流程尚属摸索阶段，还没有形成行业共识。

Pre-DevOps 阶段：ITIL，DevOps 等理念在这个时期提出，ITIL 强调流程管理质量，而 DevOps 强调打破开发，测试，运维的边界，One Culture as One Team with Closed Cycle，这个时期也开始了围绕如何落地 DevOps 工具链的技术研究，业内就 IT 研发与运维逐步达成了共识。

DevOps 阶段：DevOps 的工具链已经比较成熟，甚至出现了一些高级形式，比如 SRE，ChatOps 等，其中 ChatOps 通过对大量运维工具的封装，构建了一个代理，它认识人类定义的特定文本指令，并按照指令处理问题。这个时期自动化运维出现了，更加强调从运维流程，运维措施等层面实现完全的自动化，在特定情况下，甚至实现无人干预。

AIOps 阶段：自动化运维带来了很大进步，但毕竟系统软件是死的，只能 100% 按照人类制定的流程来运行，不能自主适应，甚至不能处理“相似”的“新”问题。于是 AI 被尝试运用到 IT 运维这个领域，这个阶段应该说才刚刚起步，行业对 AIOps 充满期待。

InfoQ：在运维过程中，有哪些技术痛点适合使用人工智能技术来解决？

张真：我认为有两类痛点可以关注：

1. 时效类问题：运维的本质是提供稳定可靠的服务，而达成这个目标的关键是足够好的时效。时效类的场景还是很多的，例如更短的 MTTR（平均故障恢复时间），特别在服务规模很大的情况下，监控数据的获取 / 集中 / 分析，问题的跟踪 / 定位，恢复的执行规划，如果再加上海量数据和状态频繁变化，AIOps 时效都会远高于有经验的人 + 工具。此外，人类有“工作时间”和“工作活力”的限制，自动化依然离不开人的决策，但智能化可以自主决策，当然目前是在经过检验的人类经验范围的扩展学习。“无中生有”的经验创造依然是个难题。

2. 协作类问题：人类的生产离不开协作。尽管有了自动化运维平台或工具链，运维很多场景还是需要许多人工协作。一个经典的例子：业务

发现了问题提交工单给 IT 服务台，IT 服务台根据经验初步判断可能与哪些系统相关，再通知相关团队，相关团队判断是否是自己的问题，如果是自己的问题则考虑修复方案，然后修复，再反馈给 IT 服务台，通知业务。这是典型的 ITIL 流程。

可是实践表明，科学的流程未必带来理想的结果。因为这个过程中人既是参与者，也是驱动者，人可能懈怠，可能 miss 信息，可能误解，可能情绪化。

另一个例子：自动化运维系统能够通过报警通知系统团队，比业务更快发现问题从而解决问题，甚至直接通过重启等自愈手段自动的解决问题，这是自动化运维带来的价值。但同样也要看到这里的问题判断与恢复规划仍然是人做的，自动化自愈等也只是人把某种情况下的问题识别，判断和处理的经验封装成执行代码而已，如果情况发生改变，系统将“不知所措”；而且系统团队也可能不了解业务影响，还是要找业务团队确认，如果业务团队太多，还是要通过 IT 服务台。那么这里的问题是什么呢？其实是缺少一个“全知”（掌握业务，系统，基础，组织的各种信息）能够客观的，全面的“协调”人，系统，业务的角色。

InfoQ：可否谈谈宜信 AIOps 探索情况？你觉得什么样的团队来搞 AIOps？

张真：首先，来谈谈背景和原则。在规划 AIOps 项目之初，我们确立了几点原则。目标是 从实际痛点入手，找到适合场景以及正确的问题来试点，而不是“大而全”的 AIOps 解决方案。技术选型上 充分利用已经比较成熟的开源 AI 技术，可以做必要改进，但尽量不重复造轮子。充分使用我们现有的 DevOps 工具链，而不是全面推倒重来。

AI 技术还不是“平民技术”，尽管已经发展了很长时间，也有人说我们处于第二次人工智能革命，但它的投入产出比可能并不像使用 Spring、Tomcat、RabbitMQ 这些开源技术栈那样的直接。所以先做“点”的事情，再考虑“面”。而且确实并不是所有场景都适合。前面也提到了，要避免凡事尽 AI。

其实问到什么样的团队来搞 AIOps。这个事与技术选型相关，也与团队定位相关。我们团队的定位是 AI FE，是将 AI 技术工程化的团队，这样的团队应该具备几个特征：

- 对现有 AI 技术充分了解和掌握。
- 选择较成熟的开源 AI 技术是必由之路。
- 对运维领域的技术（比如监控、容器技术、CI/CD、问题诊断等）是清楚的，最好是专家。
- 对运维领域的场景是熟悉的，明白运维的标准，逻辑，原则。

另外，尽管 AIOps 会带来颠覆性的运维思维和效应，但是否也要对现有系统软件来一把推倒重来呢？这里的考虑是我们的 DevOps 工具链已经比较成熟且运行稳定。同时正如前面提到 AIOps 并非是要取代现有系统，而是赋予现有系统智能。所以与 DevOps 工具链深度集成是必由之路。复用现有 IT 优良资产，最大化资产价值也是必要的考量。

再来谈谈进展。我们目前 AIOps 落地的形态是任务机器人，相关技术也围绕它展开，涉及自然语言处理、搜索技术、知识图谱、监督学习、在线学习、深度学习等。现在处于实验落地阶段，有三个基本场景。

DevOps 的一个典型场景：系统上线。上线的几个痛点是时机选择，上线条件判断，部署验证，功能验证。这些部分有的是需要人工判断的，有的通过工具进行，但都是人工驱动的逻辑判断。这是个时效类场景，如何上线更快，更可靠。

另一个场景是运维的日常工作：巡检。尽管监控系统已经可以掌握全方位的数据，比如应用性能，日志，调用链，基础设施等，还是需要有人值守；而当报警出来的时候，往往又滞后了；此外微服务架构下，人工也跟不上规模的增长和状态的快速变化。而任务机器人是可以正真全天候运行的。这也是时效类的场景，对问题的及时发现，甚至预判。特别值得一提的是，这是主动行为，而系统上线是被动触发，这两个场景正好体现了类人化智能的两面。

我们相信运维的价值在于更好的业务价值转化，Better ITOps for

Better Business。这个场景是协作类的，涵盖运维和运营。从业务同事来看他们有两个痛点：一方面他们不懂 IT 术语，玩不转运维系统，但也想时刻掌握系统运行状况；部门以及团队在运营过程的信息不对称，不能随时快速同步，造成运营效率下降。任务机器人作为中间协调者，所有人有问题就找它，它会“不厌其烦”的，“孜孜不倦”的予以解答。

如果说我们通过 AIOps 有什么收益，从前面提到的场景的痛点出发，收益是显而易见。在此次大会的分享中会对这三个场景做深度解读。

关于下一步计划，主要会考虑三个方面：

- 不断提高基本意图理解，系统 API 理解以及个性化交流语义理解的正确率。
- 加强自主问题诊断分析上的研究和应用，希望从离线方式逐步转变为在线方式。
- 尝试在更多时效类，协作类场景中应用。

InfoQ: AI 的前提是数据，那你们数据是怎么来的？

张真：关于数据来源，诚如我提到的，深度集成 DevOps 工具链是必由之路（重要的事情说三遍），因为它们就是数据的来源。当然其中监控系统是主要的数据提供者，我们的监控系统代号 UAV（含义：无人机），它提供了几种主要数据：应用画像，服务图谱，应用性能，基础设施性能，日志，调用链，业务指标。

比如通过对应用画像的学习，提取 API 模型，让系统可以使用 API，这是一种新的系统关联方式。又比如通过对服务图谱的学习，让系统掌握应用之间的关联关系，这是自主跨应用问题跟踪和影响分析的基础。还可以通过对应用性能指标的特征提取，找出异常点等。

此外，UAV 会在 9 月份正式开源，与 CNUTCon 2017 大会共襄盛举，它不但能够帮助大家实现三维一体（业务，应用性能，基础）的监控，也能如同我们一样便捷的，集中的获得 AIOps 的机器学习数据来源，欢迎大家的关注。

InfoQ: 任务机器人落地过程中，难点是什么？

张真：从我们的实践经验来看，实现任务机器人有三个主要难点：

基本意图理解：就是要理解人想做什么，这是类人交互的体现。目标是能够从自然语言中提取目标信息，并与可识别目标进行匹配，从而理解意图。我们采用了词向量与句型匹配相结合的手段，并对词向量的实现方法做了一些改进，以大幅缩减词向量空间，提高词向量的匹配速率。这个部分会介绍词向量与句型匹配相结合的基本意图理解的原理。

系统 API 理解：除了需要理解人的意图，任务机器人也同样需要像人一样去理解与之交互的系统 API 的含义以及如何使用，而且要自动适应系统 API 的变化，这是自主决策，理解执行的体现。我们采用了“微智能”（自动发现，自我维护，自动适应）与半监督学习类算法相结合的手段，让它能够认识并使用系统 API。这个部分会介绍我们的 API 模型库如何建立以及如何应用。

个性化交流上下文构建及语义理解：FreeStyle 的交流方式是不限制人必须记住某个指令或特定关键词（与 ChatOps 的区别），这是我们的基本目标。

在金融运维 / 运营的垂直领域，尽管比广义领域的词量范围要小，但仍需要解决以下问题。

由于每个人的认知差异（不同专业背景，不同的个体说话习惯），所以会用不同的词汇与句式来描述同一个事物。例如 SRE 会说“贷款系统是否健康”？应用研发会说“贷款平台有没有线上 bug”？业务同事会说“城市信贷业务运行得怎样”？实际上都是指贷款系统的服务运行情况，有没有系统或业务异常。

提取调用系统 API 的参数。NLP 只是按照人的语言习惯来提取语素信息，而任务机器人还需要从自然语言中，识别要调用的系统以及相关 API 的实际参数。例如说“告诉我电签的运行状况”，从基本意图识别来说“运行状况”指的目标服务是监控平台，“电签”是要提取的监控信息的参数。

虚拟座谈会：聊聊 AIOps 的终极价值



简单来说，AIOps 就是希望通过人工智能的方式，进一步提升运维效率，包括运维决策、故障预测和问题分析等。在 InfoQ 最近的一些文章中，都有不同程度地聊到 AIOps 相关的话题，比如美丽联合集团运维经理赵成认为 AIOps 必定是运维的发展趋势，宜信技术研发中心高级架构师张真系统分析了他们的实践案例，AliExpress 的周志伟也表示他们正在落地智能驱动的 SRE 理念。

从历史发展的角度来看，这些年，运维平台大致经历了流程化 -> 工具化 -> Web 化 -> 自动化的演进历程。随着运维管理复杂度的提升，以及企业自动化运维体系的成熟，运维平台必定会向智能化靠拢。而从结果来

看，智能化才是运维平台的最终目标。正如 InfoQ 的另外一篇文章所言，在这个数字化转型的年代，任何使用传统技术来管理机器数据的企业要么是忽略了信息的价值，要么已经让他们的运维团队不堪重负。随着数据的暴涨，运维团队应该快速拥抱 AIOps。传统 AI 仍然会在某些领域发挥它的作用，而 AIOps 将为企业带来最直接最深远的价值。

讨论组成员

- 曲显平：百度运维部技术经理
- 万金：ThoughtWorks 咨询师
- 涂彦：腾讯游戏运维总监

InfoQ：如何理解 AIOps？AIOps 里面会涉及到哪些技术？这又是一个新名词吗？

曲显平：AIOps，Gartner 有报告解释为 Algorithmic IT operations platforms，当然国内大多理解为 AI + Ops，智能化运维。涉及的技术概括来讲就是 ABC，AI + BigData + Cloud，当然 AI 的部分，不仅包括时下流行的机器学习等技术，其实也要包含传统的数据挖掘等方法。

百度从 2014 年初提出智能化运维的思路，这个名词在我们这里不算新了。

万金：AIOps 是一次跨界创新，它结合了运维场景和运维数据，使用人工智能方式试图将运维决策自动化。

AIOps 技术会涉及到数据收集方面的基础监控，服务监控和业务监控，甚至会涉及到与持续交付流水线的数据和状态整合（比如在软件发布的阶段会自动关闭某些监控项。异常判断时会参考流水线目前的状态）。数据存储与人工智能技术，其中人工智能包括机器学习算法与深度学习模型（用于模式识别）。

AIOps 可以算一个新名词，他会是比自动化运维更高级的阶段，即为了保证确定的运维目标（SLA）使用人工智能自动决策阶段。（自动运维是将重复出现的运维动作自动化，而需要人来判断什么情况（或是条件出

发) 执行那种自动化过程。)

涂彦: 不论是 AIOps 还是 OpsAI, 都是智能运维。简单来说, 就是把成熟的人工智能技术应用于互联网及互联网 + 的运维工作场景中。AIOps 可以认为是运维岗位一个分支, 本质还是运维。AIOps 本身也是由多个子岗组成的, 比如智能场景规划、数据清洗、机器学习开发等。如果从运维的发展历史来看, 确实是一个新名词。也代表了运维是个在未来有很大发展潜力与核心竞争力的工程师岗位。

InfoQ: 你认为 AIOps 是运维发展的必然趋势吗? 从手工运维, 到自动化运维, 再到现在的 AIOps, 谈谈你理解的运维发展趋势?

曲显平: 很多人认为, 运维是手工 → 自动化 → 智能化, 但其实百度不是这样认为的, 自动化运维可以说是终极状态, 不管是 Web 化、平台化、还是智能化, 用写死的 if/else 还是机器学习模型, 都是实现的方式而已。其实运维人, 要做的, 就是一切皆自动, 其实 AIOps 也就是讲的如何把人的决策也自动化起来。

为什么说他是必然趋势呢? 主要还是因为传统方法仍然不能解决运维很多问题。比如在百度, 一线运维最主要有 3 类场景: 变更管理、故障管理、服务咨询。

- 变更管理, 相比之下自动化程度最高, 但仍然不能做到完全无人值守, 最主要的问题是, 如何检查一次变更是否符合预期、变更过程中遇到问题如何处置等。
- 故障管理, 这个问题的复杂程度就比变更管理高很多了, 从应该部署什么样的监控、如何设定报警、收到报警后如何判断并做出止损操作、止损后如何做根因分析、case study、如何彻底解决这一系列问题等等, 这些还都是原有技术无法解决的。
- 服务咨询, 这个领域向来都是自动化程度最低的, 也一直是亟待解决的。

其实从这些描述来讲, 在传统意义的自动化运维过程中, 大家更多地是完成了两件事: 一个是平台化, 不再需要上到服务器去操作了; 一个是

大规模并行，一个任务可以同时成千上万个实例生效。但一旦问题变得复杂，不是 if/else 能够解决，而需要人工决策的时候，传统自动化就不灵光了。

这种需要人工决策才能解决的问题，是制约运维走向终极自动化的主要障碍，从现今的技术发展来看，AIOps 显然是最合理的路。

万金：AIOps 是运维的必然趋势，随着数字化转型的推进，业务快速扩张规模不断扩大，监控规模与数量（系统，服务和业务监控项）越来越多，云计算又导致更进一步的集中化。这些因素使得运维工作复杂度超越人力所能管理的程度，必须通过自动决策方式辅助人来进行运维工作。而对于业务而言希望不通过增加运维团队规模的方式支持迅速扩张的业务。以上两点导致智能运维的必然性。

AIOps 并不是一开始就出现的，运维的技术是随着业务的规模与复杂的提升而不断演进的，为业务选择合适的运维方式是首要考虑的问题。相比较而言，智能化运维需要投入很大的启动成本，但是能达到前所未有的运维效率；而反观手动运维启动成本很低，适合在业务规模很小的初期使用，不过云原生应用和云计算平台降低了 AIOps 的启动成本，可以帮助快速扩张的业务解决手动或自动运维无法解决的问题。

总结运维工作的特点，业务在初期都是需要手动梳理运维业务的问题，从中发现系统性问题，从而为自动化运维提供知识储备；到了自动化运维阶段会不断的标准化运维管理对象，并积累运维数据为智能运维做准备；当业务处于爆发期通过 AIOps 方式自动化保证 SLA 的执行就成了顺理成章的选择了。这里所说的无法通过自动化批量执行的任务，往往需要根据实际情况（不能通过阈值判断的情况）进行判断后采取行动。这样的任务正是 AIOps 的入手点。

比如海量阈值的自动设置来降低监控系统的错报和漏报情况，或是在海量的分布式告警信息中找到故障的根音。

涂彦：AIOps 是互联网大势所趋，也是运维发展的必然趋势，更是现阶段运维发展的终极目标，相信在未来 5 年内将更为广泛应用在不同行业

的运维工作中。自动化运维是一个承上启下的转折点。从手工到自动化，需要跨越标准化。而从自动化到智能化，需要跨越数据化。标准化和数据化是智能运维建设要思考的底层问题。从组织能力到技术能力，缺一不可。

InfoQ: AIOps 的出现是为了解决哪些问题？这些问题自动化运维没办法解决吗？

曲显平：这个问题在上一问解释过了，AIOps 的出现是为了让运维早日达成完全的自动化运维，也就是解决人工决策 → 自动化决策的问题。

万金：任何运维问题本质上都可以通过手动或自动化方式完成，但是如果考虑到 MTTR（平均故障修复时间）对业务的影响和对运维团队人数的限制，就必须引入 AIOps。

首先，运维每天最多的工作就是监控系统是否正常运行，如果出问题了就需要及时解决，比如 99.99% 的可用性就意味着每年只能出现 52.56 分钟的系统不正常时间，一般人工处理一个故障在熟练的情况下也会在 20 分钟也就是说在频繁上线更新的情况下只能期望一年系统不要出 2 次以上的问题。如果想提高系统可靠性就必须引入 AIOps。

其次，AIOps 的出现大大缩短了 MTTR，在解决如何发现系统异常（自动化监控项管理）和如何找到问题根因（告警抑制：通过关联性判断，把计算资源、软件和网络具有相关性的告警信息聚合，让运维人员迅速找出问题根因的技术）。

最后，人的管理能力是有限的，就像人可以驾驶汽车，但飞机就需要自动导航只在启动和降落需要人工干预，对于宇宙飞船，宇航员只有在出现故障的时候才会手工干预。AIOps 就像驾驶宇宙飞船的计算机，只有 AIOps 的自动决策结果与预先设定好的目标（SLA）不相符才需要人工干预。

涂彦：质量、效率、成本、安全，是运维工作核心四要素。用 AIOps 来解决，会在提升效率的同时，将质量与成本更加精细化，使安全的应变能力更强。自动化在决策方面很难快速适应千变万化的生产环境和业务需求，运维经验仍然是主导地位，依靠个人能力的痕迹明显，对企业管理的风险都受控于个人。

InfoQ：落地 AIOps 的前提条件是什么？什么样的团队适合落地 AIOps ？

曲显平：从百度运维的 AIOps 经验来看，落地 AIOps 的前提条件是已经具备了比较完善的运维平台和有较充分的运维数据，也就是 ABC（AI、BigData、Cloud）里的 A，离不开 B 和 C 的支撑。

我们的团队，用了 5-6 年的时间才完成了运维的完全平台化，以及建立了统一的运维数据仓库，在这之前当然也可以落地一部分 AIOps 场景，但相对应的，也是需要这部分场景的 B 和 C 具备。目前看，应该是运维平台和数据都已具备时，才是落地 AIOps 的理想时刻。

万金：在那些具有较强自动化运维能力，和一定的数据储备的条件的团队才适合 AIOps 落地，同时业务是否对运维效率提升有需求也是一个考虑因素。

引入 AIOps 之后，是否能对 AIOps 的模型或数据进行不断优化也是一个新的挑战。人工智能在一开始都不是很完美的，需要不断优化才能达到实际应用的要求。对于发现问题和问题根因分析方面的 AIOps 落地速度比较快，对于高级阶段的根据 SLA 自动调度的 AIOps 就需要比较长的优化

涂彦：标准化、工具化、自动化、数据化、场景化是前提条件，同时这也是逐一递进的关系。

随着人工智能技术的愈加成熟，只要具备业务需求的团队，都适合落地。人工智能技术在运维行业的应用与其在其它行业落地的状况类似。所以说，AI 并不是那么神秘，使用的门槛只会越来越低。但是如果要想在企业中真正用好，需要对业务理解并紧密结合技术方案的运维，比如通过智能场景规划来找到痛点，再结合数据清洗与机器学习开发来完成落地。

InfoQ：AIOps 中的数据是怎么来的？数据是必要的吗？

曲显平：数据非常必要，因为 AIOps 中的 A 主要指的是算法（数据挖掘、机器学习），每一个算法都离不开数据的支撑，尤其对于深度学习而言，普通体量的数据都不足以支撑训练出理想的模型，需要非常庞大的数据量。

AIOps 的数据，从来源上看，主要分四类：设备、系统、平台、业务，设备主要指 IDC、网络、服务器等偏硬件、数据中心层的设备信息；系统主要指操作系统等；平台主要指基础运维平台、PaaS 平台等；业务主要指产品服务的日志等。

AIOps 的数据，从类型来看，主要也是四类：一个是时序数据，这个重要性最高，因为其结构化层次高，在百度，这个数量级能达到十亿级。第二个是运维事件数据，这个同样也非常重要，每一次异常事件、变更事件、运营事件等等，都是需要被记录并加以分析的。第三个是日志数据，相比之下日志数据的结构化偏弱一些，但同样十分重要，由于量级太过庞大，我们也只会挑选并存储较为重要的部分，来进行分析训练。当然，还有最后一类，因 AIOps 而产生的数据“标注数据”，这部分数据的完善程度将直接影响着每一个算法模型的实际效果。

万金：AIOps 中的数据是从监控系统和相关的运维经验中来。而 AIOps 也是有不同的实现的技术，比如通过机器学习算法的 AIOps 依赖数据较小，只要找到合适的算法就能对当前数据进行处理，但是处理效果随着数据模式的改变而改变，也就是当数据体现的模式改变后就必须手动更换算法适应新模式。而深度学习的算法就不必人工更换模型只需要设定目标（SLA）但对数据依赖比较多。

涂彦：训练的基础就是数据，这个与其它行业并无区别。数据来自于运维工作与服务的场景。质量、效率、成本、安全，任何一个维度都需要将数据标准化采集，这是一切的基础。再比如技术运营的一些触达到用户或者产品的场景，也是一个数据化的过程，任何一个节点的数据都可以被采集。这些标准化和数据化的工作，都是 AIOps 的基础工作。

InfoQ：可否谈谈你们的 AIOps 落地场景？

曲显平：在百度运维，AIOps 应用的很广泛，监控异常检测、故障诊断分析、智能流量调度、SQL 入侵检测、成本优化、性能分析优化等等。

我们在自动化异常检测方向的研究非常早，早期还以传统的时序数据分析为主，现在由于机器学习等方法的兴起，已经多样化很多，在百度大

多数核心时序指标的监控都是用的自动化异常检测模式。

在故障诊断方向，这一直是一个比较难的课题，我们研究过非常多的行业 paper，真正实现后效果出色的少之又少，当然这也可以理解，因为对于人来讲，诊断故障也一直是一个让人头疼的问题。

此外，在智能流量调度、SQL 入侵检测、成本优化、性能分析优化等层面，我们应用了非常多的机器学习模型，为公司业务的可用性、成本、性能等的提升做出了巨大贡献。

万金：故障识别和故障自愈，资源自动调度。

更高级的场景会引入全自动软件发布控制与回滚，自然语言容量管理，主动安全识别和在软件研发完成前对软件体验进行评分。

涂彦：腾讯游戏的运维团队聚焦在基础服务与增值服务中落地。基础服务包括发布变更、故障管理、用户体验服务。增值服务包括触达到用户与产品决策的运维扩展服务。发布变更中的数据预测、故障管理中的根因分析、用户体验中的多维告警等，都是 AIOps 的落地场景。触达用户的云控策略、产品决策的舆情分析等，也都有着很好的应用场景。

InfoQ：目前业界有哪些可以参考的 AIOps 实践？

曲显平：在 AIOps 和智能化运维这个角度，百度运维是做的比较早的，我们将运维能力标准定义为六级，L0-L5，每一级都有相对明确的定义，很多方向都可以应用 AIOps 来助力等级的提升。

当然除了百度，推荐看下 Netflix 的 Winston，以及 Google 的 Auxon，这些实践，不仅是用某种算法在解决一个问题，他们正在尝试构建一种开放的自动化 / 智能化运维模式。

越来越多的公司关注并且重视这部分信息的开放和共享，从场景到算法，从数据到开放框架，百度运维也正在参与其中，也欢迎大家合作共建智能化新运维。

万金：在我的文章里提到一些算法方式的实践，深度学习方式由于依赖多，成本高不确定性大应用比较少。行业参考如下：

- Twitter: Seasonal Hybrid ESD (S-H-ESD)

- Netflix: Robust PCA
- LinkedIn: exponential smoothing
- Uber: multivariate non-linear model

涂彦：腾讯蓝鲸智云来自于腾讯游戏数百款游戏的基础服务与增值服务的最佳实践。目前已经对外提供社区版和企业版。蓝鲸可以帮助运维团队在标准化、工具化、自动化的深度建设，结合蓝鲸数据平台，AIOps 将提供给运维团队更多想像空间。

版权声明

InfoQ 中文站出品

深入浅出 AIOps

©2017 北京极客邦科技有限公司

本书版权为北京极客邦科技有限公司所有，未经出版者预先的书面许可，不得以任何方式复制或抄袭本书的任何部分，本书任何部分不得用于再印刷，存储于可重复使用的系统，或者以任何方式进行电子、机械、复印和录制等形式传播。

本书提到的公司产品或者使用到的商标为产品公司所有。

如果读者要了解具体的商标和注册信息，应该联系相应的公司。

出版：北京极客邦科技有限公司

北京市朝阳区洛娃大厦 C 座 1607

欢迎共同参与 InfoQ 中文站的内容建设工作，包括原创投稿和翻译，请联系 editors@cn.infoq.com。

网 址： www.infoq.com.cn