IBM Security

安全和人工智能

常见问答

Jeff Crume
Doug Lhotka
Carma Austin

安全和人工智能：常见问答

引言

就网络安全而言，我们都知道：变化是永恒的。因此，我们必须不断审视过往，识别其中的改善方式。若要跟上网络攻击者的步伐，我们必须不断尝试新技术，找出防范或主动预防攻击的更好方式。我们必须每天评估我们的策略并提升我们的方法。简言之，如果我们不作出改善，我们就无法确保最重要资产的安全。

每个安全供应商都了解"变化是永恒的"这一基本原则。几乎每年在网络安全领域都会出现几个新的热门词，说明我们一直都在找应对安全挑战的解决方案，但让我们感到惭愧的是，收效并不理想。在最近几年里，我们的收件箱里充斥着"可执行"和"自动化"等字眼，而现在，我们又看到了新的热门词，包括人工智能 (AI)、认知计算和机器学习等等。

如今，许多供应商都表示 AI 是他们的秘密武器，事实也的确如此。这种关键的技术进步似乎是一夜之间就完成了"蜕变"，但许多行业专业人士对 AI 的前景仍旧充满了怀疑。可惜的是，AI 流行的同时也带来了混乱，因此妨碍了整个市场真正地采用该项技术。就目前而言，"AI 是否存在？"，这个问题的答案已经显而易见。因此，我们要关注的问题是"什么是 AI？"以及"AI 能为我们提供哪些帮助？"。这些就是本文的目的所在：抛开诉求和愿望不谈，而是定义出 AI 及其组件能够为当今的安全领域提供哪些帮助。

常见问题

- 什么是人工智能？

  人工智能 (AI) 在《韦伯斯特辞典》中的定义是："人工智能是计算机科学的一个分支，旨在解决在计算机中模拟智能行为这一问题。"[1] 《牛津字典》中也给出了 AI 的详细解释，即"人工智能是指通过计算机系统的理论和开发来执行通常需要人类智能才能完成的任务，比如视觉感知、语音识别、决策、语言翻译等等。"

- 这些定义都切中了要害，但如果简单来说的话，什么是人工智能 (AI)？

  从本质上来说，人工智能是指让计算机模仿人类进行发现、推断和推理，实现与人类智能相当或超越人类智能的能力。

[1] Merriam-Webster，https://www.merriam-webster.com/dictionary/artificial%20intelligence

[2] 牛津字典，https://en.oxforddictionaries.com/definition/artificial_intelligence

安全和人工智能：常见问答

- 我们是否可以说，AI 并非另一个营销热门词？它是真实存在的吗？

  许多事情都会借助 AI 来大肆宣传，导致扭曲了 AI 的定义。换句话说，有关 AI 的一些承诺已经变成了现实，但其他一些仍旧处于研究阶段。与其说 AI 是一个单体功能，不如说它是支持同一目标的许多相关技术的合集。举例来说，以下所列的各种功能如果结合使用的话，就有助于让计算机变得更加智能：

  - 推理和问题解决
  - 知识呈现
  - 规划
  - 学习
  - 自然语言处理
  - 感知
  - 机动性与自主操控
  - 社交智能
  - 创意
  - 通用智能[3]

- 什么是机器学习？

  机器学习是 AI 和计算机科学的一个子领域，其根本在于统计与数学优化。机器学习包含了预测应用、分析应用和数据挖掘应用所使用的监督式学习与非监督监督式学习中的技术。[4] 机器学习可以独立于其他 AI 或认知技术来使用，而且通常的情况正是如此。事实上，机器学习是我们目前最常见的一种 AI 类型。许多机器学习算法和技术已经在各种解决方案中得到了运用，这些解决方案的目的在于寻找数据中的范例或异常。

- 什么是深度学习？

  深度学习是一种相对较新颖的方法集，它正在从根本上改变着机器学习。从本质上来说，深度学习并非一种算法，而是指通过非监督式学习实施深度网络的一系列算法。这些网络的深度非常之大，因此需要采用 GPU 等新的计算方法才能得以构建。[4]

---

[3] 维基百科，https://en.wikipedia.org/wiki/Artificial_intelligence

[4] M. Tim Jones."A Beginner's Guide to Artificial Intelligence, Machine Learning and Cognitive Computing". 2017 年 6 月. https://www.ibm.com/developerworks/library/cc-beginner-guide-machine-learning-ai-cognitive/cc-beginner-guide-machine-learning-ai-cognitive-pdf.pdf

安全和人工智能：常见问答

- 深度学习听起来非常不错。就它的运用而言，是否存在局限性？

  没错。尽管深度学习算法的运用取得了一定的成效，但仍旧存在许多我们尚无法解决的问题。就近期深度学习在皮肤癌检测方面的运用效果来看，这种算法的确比经过认证的专业医师更加准确。但专业医师能够列举出他们作出诊断的依据，而深度学习程序在进行分类时所用的依据，我们却无法识别。这就是所谓的深度学习的"黑盒子"问题，这就给模型确认带来了诸多挑战，尤其是在获得监管许可方面。相关证据都显示，这种算法很准确而且有效，但如果我们无法描述它的决策机制，对它的完全确认也就无从谈起。

  有一个应用 Deep Patient，它能够依据病患的医疗记录成功地预测疾病。经过验证，该应用在疾病预测方面要远远优于医师，甚至是对于普遍认为难以预测的精神分裂症而言，该应用也能成功预测。因此，尽管深度学习模型的效果很好，但没有人能够深入到其中的广泛神经网络找到背后的原因所在。[4]

  深度学习在特定问题的解决方案具有很大的潜力，但并不适用于所有情况，而且实施起来非常困难，不仅成本高而且非常耗时；只有在很少一部分情况下，它才能够发挥最大效力。

- 什么是认知计算？

  认知计算是 AI 的子领域之一，它基于神经网络和深度学习而构建。它采用来自认知科学的知识来构建能够模拟人类思考过程的系统。不过，认知计算并非仅仅专注于单个技术集，而是涵盖了多个学科，包括机器学习、自然语言处理、视觉和人机交互等等。[4] 当然，认知计算的重点在于自然语言处理。

- 认知计算为何如此特别？

  以下内容节选自 IBM 研究总监 John Kelly 撰写的一篇白皮书[5]，其中很好地解答了这个问题：

  认知计算系统是指能够大规模学习、进行目的性推理，并自然地与人类互动的系统。无需进行明确的编程，认知系统便可了解人类并根据与人类的互动和环境体验进行推理。之所以能够实现这些功能，得益于超过半个世纪以来许多科学领域的研究进展，而且认知计算在很多方面都与之前的信息系统完全不同。之前的系统是确定性的，而认知系统是概率性的。认知系统不仅能生成许多问题的答案，而且会针对更复杂且更具意义的数据主体生成假设事项、推理得出的论点和建议。此外，认知系统可以利用全世界 80% 的数据，计算机科学家将这部分数据称为"非结构化数据"。正因为如此，认知系统才能在数量、复杂程度和不可预测性方面跟上当今时代的信息和系统的步伐。

---

[5] Dr. John E. Kelly III." Computing, Cognition and the Future of Knowing: How Humans and Machines are Forging a New Age of Understanding". 2015 年 10 月. https://cra.org/crn/2016/09/computing-cognition-future-knowing-humans-machines-forging-new-age-understanding/

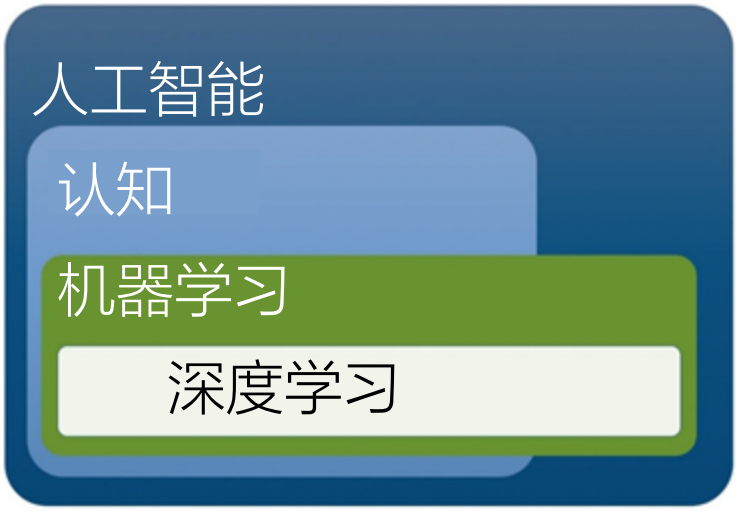
安全和人工智能：常见问答

- 认知计算是否与 AI 相同？

  认知计算是 AI 技术的一种，但也有一些 AI 技术与认知全无关系。

- 所有这些 AI 技术之间有何关联？

  深度学习是机器学习的一种特殊形式。认识计算会同时使用这两种技术。所有这三种技术都属于 AI 技术。下文的文氏图就很好地说明了它们之间的关系。



- 什么是 IBM Watson® ？

  Watson 是 IBM 推出的 AI、机器学习和认知计算平台。该平台提供了一系列 AI 技术，可处理来自各种来源的结构化信息和非结构化信息，理解这些信息的涵义，然后将其添加到它的知识体系（即知识库）之中，以供后续使用。之后，该平台的一些组件能够使用基于该知识库的自然语言来响应复杂的问题，既可以增强人类智能，还能够改善关键任务中的洞察力和效率。其他组件可提供机器学习和深度学习功能，以供在处理结构化数据时使用。

  Watson 平台的解决方案并非单个产品，而是一整套集成式的组件，它们具有专门的分析功能，以供各个组件，以构建支持 AI 的解决方案。尽管所有的组件都与 AI 有关，但一些组件还可提供认知功能，许多组件还使用了机器学习或深度学习，还有一些使用的是更加传统的分析功能。

安全和人工智能：常见问答

- 听起来非常不错，但 AI 和认知计算是否可用于解决实际问题？

  没错。举例来说，北卡罗莱纳大学教堂山癌症中心已经开始使用 IBM Watson 为常规治疗无法见效的病患识别和推荐治疗选项。每天医疗界大约会发布 8,000 篇新的医疗研究文章，因此任何一个医师，甚至是一个医学团队都无法跟上最新研究成果的步伐。但 Watson 可以利用所有这些最新研究成本，然后快速地加以运用，用于改善病患治疗成效。在针对 1,000 名病患进行的测试中，Watson 的治疗建议与医疗专家所给建议的匹配率高达 99%。更重要的是，在专家都未察觉的病例中，有大约 30% 的病例 Watson 可以发现其他治疗选项。[6]

  这仅仅是许多 Watson 运用示例的其中之一。 另一个示例是，IBM 已经开始将 AI 和认知技术运用到网络安全领域，旨在帮助组织更快速地识别和响应威胁。Watson for Cyber Security 已经在知识库中摄入了超过 20 亿篇文档，而且每天都会增加数千个新文档。它能够将意外事件分析所需时间从数小时缩短到数分钟，极大地加快了意外事件缓解速度，也减少了意外事件对组织的影响。

- AI 是否会淘汰人类？

  不会。AI 的目的是增强人类智能，而非取代。认知计算的用途仍旧存在很大的局限性，尤其是在决策领域，因为在该领域，许多人类可以权衡的因素无法轻松地通过算法术语来表达。AI 仅仅是一个工具。与其他工具一样，AI 也能够增强人类所从事的工作。AI 能够从整体上处理海量的非结构化数据，因此能够显著提升效率和洞察力。

- AI 是否会变得有意识进而控制地球？

  我觉得您可能是科幻片看多了。

- AI 是否是“黑老大”？

  与任何技术一样，AI 也可能会被误用。举例来说，Watson for Cyber Security 需要使用数据加以训练，才能提供识别威胁和攻击场景所需的洞察力。向 Watson for Cyber Security 发送何种类型的数据进行分析，配置的控制权完全在您手中。我们都知道，攻击者非常喜欢通过滥用 AI 来实现他们自身的目的，但就目前而言，我们在 AI 使用方面要远远领先于攻击者。AI 仅仅是一个工具，因此既可以用于好用途，也可以用于坏用途。

---

[6] “Artificial Intelligence Positioned to Be a Game-Changer”. *60 Minutes.* CBS News. 2016 年 10 月 9 日. https://www.cbsnews.com/amp/news/60-minutes-artificial-intelligence-charlie-rose-robot-sophia/

安全和人工智能：常见问答

- 什么是预测性分析？

  顾名思义，预测性分析是高级分析的一个分支，主要用于预测未知的未来事件。预测性分析能够通过数据发掘、统计、建模、机器学习和人工智能等多种技术从数据中提取信息，然后分析当前数据，做出未来预测。

- 预测分析能够为我们提供哪些帮助？它有哪些局限性？

  关于这个问题，我觉得应"视情况而定"。分析的预测能力取决于所用数据集的质量和相关性，以及处理数据所用算法的准确性，甚至还与负责解读从数据中所获洞察力的个人的能力有关。换言之，一些事件会比其他事件更容易预测，而且一些首次发生的事件可能永远无法预测（我们有时候会将该类事件称为"黑天鹅事件"）。

  举例来说，在安全领域，零日攻击非常难以预测，因为恶意软件的广泛散播实在是太容易了。换句话说，即便这些技术被冠以"预测性"的标签，但实际上他们提供的仅仅是未来发生某一事件的*概率*，而且从定义上来说，黑天鹅事件是完全无法预测的。不幸的是，安全领域中的黑天鹅事件并不少见。加上安全领域的预测分析用例仍属新鲜事物，因此预测结果可能会千变万化，我们不能对它抱有过高的期望。

- 什么是威胁捕获 (Threat Hunting)？

  在网络安全领域，威胁捕获是指通过搜索大量数据来识别针对组织 IT 基础架构的攻击者和威胁。其目的在于在攻击发生之前防范攻击，并消除或最大程度地减少攻击的影响。威胁捕获工具可以摄入威胁情报摘要、漏洞分析报告、风险评估结果、恶意软件分析结果、HR 员工记录、安全事件数据、系统日志、社交媒体摘要等各种数据。

  尽管威胁捕获会利用许多工具，但许多工作仍旧是由事件调查人员手动完成的，他们需要研究他们主动列出的一些问题的答案。举例来说，他们需要负责搜寻某个组织内曾访问高度敏感资源的员工，或者是近期曾对组织表达过不满情绪的员工，因为这些员工都可能带来潜在的内部人员威胁。当然，此类活动也存在一些局限性，因为它需要结合大量因素才能检测到异常活动，然后预测未来的事件。由于此类活动需要大量的人工参与，因此对它的期望不要过高。没有任何一种 AI 技术能够统一、可靠地从一片数据汪洋中找到所期望的那块"小石子"。

[7] "What Is Predictive Analytics?"Predictive Analytics Today.
https://www.predictiveanalyticstoday.com/what-is-predictive-analytics/

安全和人工智能：常见问答

IBM Security 与 AI

- IBM 推出了哪些 AI 产品？

  Watson for Cyber Security 是 IBM 认知计算功能的一个实例，主要针对的是网络安全领域。它能够使用来自威胁情报摘要、安全事件及相关数据的安全信息，也可以使用研究报告、安全博客、网站和咨询报告等非结构化来源的数据。之后，它会将这些信息存储为一个广泛的知识库，该知识库由超过 100 亿个要素构成，而且会以每小时超过 400 万个要素的速度刷新它的理解结果。从某种意义上来说，Watson for Cyber Security 就像是一个特殊的安全专家，它能够全天候读取 Web 数据，永远不会忘记，还能够根据这个极具动态性的知识库就攻击作出假设，而且随着时间的推移，会变得越来越智能。现在，您可以通过 QRadar Advisor with Watson 来访问 Watson for Cyber Security 工具。

  QRadar Advisor 是 IBM QRadar Security Information and Event Management (SIEM) 平台的延伸应用，它能够将威胁指示器发送到 Watson for Cyber Security 中进行处理，而 Watson for Cyber Security 之后会返回与网络攻击事件的性质和程度相关的洞察力。该应用能够在 QRadar（其本身使用了机器学习等 AI 技术）中合并安全分析结果，以发现异常，同时还会结合使用 Watson for Cyber Security（主要以认知技术为基础）的推理技能和知识库，以快速分析攻击事件的完全范围和情境信息。

  i2 Enterprise Insight Analysis 可用于执行威胁捕获和调查。i2 QRadar Offense Investigator 将 QRadar SIEM 的功能与 i2 集成一体，有助于改善调查效率。

- IBM 的安全产品组合中还有哪些产品也利用了 AI 技术？

  我们已开始在整个产品组合中融入 AI 功能。举例来说，多年以来，IBM QRadar 一直采用机器学习来识别环境中的安全异常事件。最近，QRadar User Behavior Analytics 应用也开始利用高级机器学习来检测用户行为方面的变化并给出风险评分。

  IBM MaaS360 采用 Watson 系列中的其他 AI 技术从非结构化数据中发掘与移动终端设备漏洞相关的洞察力。

  IBM AppScan Source 利用机器学习来减少在源代码中查找潜在安全漏洞时的误报数量。

- 真的是这样吗？是不是任何企业都可以借助认知功能来改善他们的安全态势？

  没错。使用 QRadar Advisor with Watson 的 IBM 客户通常都会发现，他们的安全态势得到了明显改善，因为借助该工具，他们能够更快、更彻底、更统一地完成调查。他们还能够高效处理他们的日常积存的事件，即便是经验较少的小规模分析师团队也能做到这一点。

安全和人工智能：常见问答

以下所列是一些真实客户的真实反馈。

智能	速度	准确性
它是市场上唯一一款能够从 Watson 所摄入的超过 100 万篇安全文档中提取洞察力，进而为用户提供与攻击相关的完全情境信息和范围的解决方案。 具有超过 100 亿个安全相关节点，可以轻松关联容易被安全分析师遗漏的隐藏威胁 *"…一级和二级分析师都认为这不是一个安全意外事件。但 Watson 给出的调查结果更具建设性。它在数分钟内就完成了定性分析，认为我们客户的某个主机遭到了 DDoS 攻击。*	比手动威胁调查快 60 倍 加快了复杂分析的速度，将分析事件从 1 小时缩短到 1 分钟以内 *"Watson for Cyber Security 能够将分析流程的速度提升 50%。如此一来，我们的员工便可在较短的时间内分析更多的信息，同时还可以针对最难缠的威胁快速采取措施。"* *"Watson 的每次分析仅需不到 1 分钟的时间，而人工分析的话，一般需要 15 分钟到 1 小时的时间。"*	发现新威胁所需可执行洞察力（指示器）的数量增加了 10 倍 无需额外的软件，也不需要深入的分析师专业知识，而且永不休假 *"QRadar 针对某个试图连接到僵尸网络 IP 的用户攻击实施了拦截。安全分析师以手动方式发现了 5 个相关联的指示器，而 Watson 则给出了该威胁的严重程度分析结果以及 50 多个有用的指示器。"* *"随着时间的推移，分析师会由于注意力分散而出错，但 Watson 永远不会出错。"*

- 安全响应是否可实现自动化？

  也是也不是。举例来说，恶意软件攻击的响应流程早在多年以前就已经实现了自动化，而且效果也非常不错。一般来说，该自动化流程是将受感染的文件移动到专门的隔离容器中，以防它们进一步散播。尽管如此，也经常会出现这样的情况：签名机制将某个操作系统或其他应用组件判定为恶意软件，进而造成系统中断。

  知名安全专家 Bruce Schneier 曾这样说道：

  > *意外事件响应从根本上来说具有不确定性，因此难以实现自动化。我之所以这么说，是有很多原因的。所有的攻击都互不相同。所有的网络都互不相同。安全环境也互不相同。组织也互不相同。监管环境同样互不相同。在多数情况下，政治因素和经济因素要比技术因素重要的多。*[8]

  从根本上来说，自动化需要确定性，而在安全领域唯一能够确定的就是多变性。在负面风险较小的特殊情况下（例如隔离某个用户的机器），可能会比较适宜采用自动化；而在业务中断风险较高的用例中（例如在供电行业或医疗保健行业），应谨慎地使用自动化。当前的最佳实践是委派一名人员参与到流程中，负责监控其中的所有关键情况。基于上述原因，我们将这种更广泛的概念称为"编排"，也有的场景里叫做"统筹"，而不是"自动化"，具体来说就是，人类负责决策环节，而机器负责快速执行已批准的任务。

[8] Bruce Schneier."Security and Privacy in a Hyper-Connected World". InterConnect 2017.

安全和人工智能：常见问答

- 有关本文所述的主题，我还可以从哪里了解到更多信息？

  以下所列为一些非常不错的延伸阅读文档：

  IBM Cognitive Security White paper

  A Beginner's Guide to Artificial Intelligence, Machine Learning and Cognitive Computing

  5 Things You Need to Know About AI Buzzwords:Cognitive, Neural and Deep, Oh My!

结论

早在上世纪 50 年代，人工智能的概念就已经提出。一直以来，该技术的目的都是通过模仿人类大脑在识别和解读复杂范例时所采用的方式，模拟能够在某一天开始自主思考的神经网络。自从上世纪 50 年代以来，我们在人工智能领域已经实现了飞跃式的发展，尽管我们永远也无法实现机器像人类一样思考的科幻景象，但我们已经迈入了认知计算时代，人机交互已经变为现实。

认知系统具有理解、推理和学习能力，而且能够以自然语言与人类交互，因此说它已经偏离了核心机器学习和机器智能的范畴。认知系统的出现意味着，我们第一次实现了计算机可以从非结构化数据中发掘我们人类也无法识别的洞察力这一目标。

IBM Watson for Cyber Security 是一个由许多不同的 AI 子领域构成的 AI 系统，包括自然语言识别与处理、预测性分析、数据挖掘、机器学习、深度学习，以及用于显示数据集之间相互关系的知识图等等。这种架构通过"合力"构建了一个广泛的知识库，其中包含有数十亿个在之前无论是人类还是机器都无法规模化访问的数据要素。通过每次体验，该知识库会变得更大、更智能，让智能信息变得触手可及。

QRadar Advisor with Watson 在非结构化数据世界 (Watson)、结构化数据的深度分析和机器学习 (QRadar) 及人类安全分析师之间搭建了一个桥梁，有助于提升安全事件调查的幅度、可视性和速度，但也仅仅是触及到了 Watson 功能的表面而已。我们无法确知未来会怎样，但毫无疑问的是，AI 时代已经来领，而且有望成为网络安全领域的"游戏规则改变者"。

安全和人工智能：常见问答

如果您希望了解 IBM 在人工智能方面的更多观点，或者希望了解有关 IBM 人工智能产品的更多信息，可以与任何一位作者联系。

Jeff Crume，杰出工程师、IT 安全架构师、IBM 杰出发明人、IBM Security 解决方案技术销售专家

crume@us.ibm.com

Doug Lhotka，网络安全执行架构师、CISSP-ISSAP

dlhotka@us.ibm.com

Carma Austin，安全执行顾问

caaustin@us.ibm.com