

Benno Keller

日内瓦协会电子及创新领域特邀顾问

在保险行业应用人工智能，可以通过提高风险共担能力，加强降低、缓解和预防风险等能力，有望产生巨大的经济及社会效益，惠及除保险公司、投保客户以外更多的相关方。为了促使人工智能系统尽早落地应用，使其实现其经济社会效益，保险公司需通过切实负责地使用新技术来赢得客户的信任。负责任的人工智能系统应体现诸如透明性、可解释性、公平、安全、负责和保护隐私等核心原则。为了普及推广这些原则，保险公司建立内部的指导方针和政策，搭建适当的治理架构以应对相关风险，同时为职工和代理人研发出一整套全面的培训项目。

### 保险行业的人工智能

“人工智能的寒冬”始于 20 世纪 70 年代末，终于 20 世纪 90 年代末，在此期间人工智能的发展之路充满挫折和沮丧。自那以后，人工智能技术取得了显著的进步。今天，许多保险公司推出了智能系统，实现了常规任务自动化，或辅助人们在整个保险价值链上做出决策。这类系统整合了新型的学习算法与来自网络媒体和物联网（IoT）等新信息源的数据分析相结合。

例如，自然语言处理的进步使得智能系统能够与人类进行“交谈”和互动。保险公司越来越多的使用会话智能体（如聊天机器人）等，这些工具能全天候地识别和响应负责的客户询问。而且，智能系统能够“查看”及辨识出图片中的对象，并提取出相关信息。通过运用这类计算机视觉技术，保险公司能使一些常规的手动和认知工作实现自动化，例如从书面文件和图片中提取数据用于承保或索赔流程。

智能系统尤其擅长在一堆复杂数据中识别出其特征和关联性，这是人类很难甚至不可能完成的任务。这些被识别出的特征是分类、回归、聚合等分析工作的基础，而这些分析工作又在保险业务模型中起着重要作用。与保险行业传统的建模方式（如广义线性模型）相比，智能系统能够掌握变量间复杂的非线性关系，因而能够提供更为准确的预测。随着学习算法的进一步发展，这些智能系统有望在更广的领域，在人的监督下自动地做出标准化的决策。

### 在保险行业运用人工智能的优点

为了促使人工智能系统尽早落地应用，使其实现其效益，保险公司需通过切实负责地使用新技术来赢得客户的信任。人工智能也能帮助保险公司增强其研敌，缓解和预防风险的能力（详见表 1）。

**表1：在保险行业运用人工智能的社会经济效益**

<b>扩大风险共担的范围</b>
<ul style="list-style-type: none"><li>· 通过提供个性化的产品（如为已有疾病的个人提供人寿保险），将保险的覆盖范围扩大到新的或之前未参保的客户群体</li><li>· 通过提高对风险的洞察力，扩大可承保风险类别（例如网络风险）</li></ul>
<b>降低风险共担的成本</b>
<ul style="list-style-type: none"><li>· 通过将特定任务自动化，实行更完善的风险评估，减少道德风险和逆向选择，为客户提供更经济划算的保险</li></ul>
<b>预防并降低风险</b>
<ul style="list-style-type: none"><li>· 培养新的风险洞察力，帮助预防并降低风险</li><li>· 建立早期预警系统减少损失</li></ul>

来源：日内瓦协会

### 负责任的人工智能系统的运行原则

过去几年，人们就负责地使用人工智能的必要性开展了激烈而持久的辩论，各种政府和非政府组织为人工智能的合理合规使用发布了诸多指导方针。对这些指导方针进行分析后不难发现，全球范围内正在朝着负责任的人工智能的五大核心原则趋同：（1）透明性和可解释性，（2）公平性，（3）安全性，（4）可说明性和（5）隐私性。然而，关于道德原则和指导方针应如何在特定情

境下实施仍存在诸多不确定性。

在保险行业，以上这些原则长久以来一直发挥着重要作用。事实上，各类法律法规，包括保险法、隐私和数据保护法、反歧视法和其他监管规定都在约束保险公司，使其行为公平、透明、负责的同时，也在保护着客户的隐私。

但是，使用人工智能也对保险公司提出了一些复杂棘手的问题。如，核心原则的实施会带来哪些利益的权衡？保险公司如何能更好贯彻这些原则？如何向客户及监管机构证明已贯彻了这些原则？为了贯彻这些原则，需要对现行的治理机制和风险管理架构做出哪些改变？越来越多的保险行业监管机构也在考虑以上问题。

日内瓦协会发布的报告《在保险行业推广运用负责人的人工智能》，分析了人工智能使用方面的一些重要道德准则，并探索了，鉴于计算机科学公平机器学习领域的发展现状，如何在保险行业贯彻这些准则。这些问题尚无确切答案。本报告的目的是为了讨论当保险行业抱着负责的态度使用人工智能技术，贯彻以上核心原则时会舍弃哪些其他关键性的收益。虽然以上原则都很重要，本报告将着重关注以下原则，即1）透明性和可解释性，2）公平性，因为通常这两类原则在保险行业会引发尤为复杂的问题。



## 透明性和可解释性

透明性及可解释性对于建立与客户和其他利益相关方之间的信任十分重要。在一些指导方针中均提到了透明及可释的重要性，为的是帮助个人就造成侵害的决策寻求赔偿。当一项决策对受害个人有显著影响时，为个人提供合理解释就变得极其重要了。因此，需要解释的程度主要取决于一个错误或不准确输出的内容和所造成结果的严重程度。算法结果的可读程度同样也是评估人工智能系统表现、系统持续改进水平，以及数据科学完备程度必不可少的先决条件。

然而，提供有意义的解释是一项艰巨的挑战，因为一些“黑匣子”算法本身极其复杂，难以解释，这也是为实现精确性必须付出的代价。近些年，人们向计算机科学倾注了大量的精力，以期战胜挑战，解释“黑匣子”算法。例如，*逆向工程法*就试图寻找建立一些可读的算法代理<sup>1</sup>，我们需要更好地理解这一新技术。*设计方法依赖*对预测施加特定的约束。

当无法解释某一特定决策中不同变量的作用时，可以使用其他方法，如由第三方独立机构为人工智能系统提供认证等，也可以用作增强相互信任。即便如此，我们仍然鼓励保险公司建立那些可读可释的模型，尤其当模型的产出结果对客户有显著影响时。当人工智能系统被用作风险选择和定价时，选择使用那些与参保风险相关的数据源，以一种客户可以直观理解的方式来增强客户对人工智能系统的信任。过于复杂的模型有时确实会带来一些好处，与此同时可读性变差也是不争的事实。

## 公平性

一项公平的决策往往存在若干相互排斥的定义，如何保证公平性尤为错综复杂。公平性常与诸如自由、尊严、自主、隐私、非歧视、平等，多样性等许多不同价值观联系在一起。这些价值观需要放在特定情境，如不同的文化情境下进行解读。因此我们无法为公平性提供一个通用的衡量标准。

决策应保证公平，保证不歧视、不剥夺个人或集体应有的权利。在人工智能主导的决策中消除或最小化不公平的偏见便成为道德指导原则中的一项核心要求。

人类或多或少都带有这样或那样的偏见。在某些情况下，运用人工智能可以使决策变得更加公平。但是，假使人工智能系统实现大规模运用，系统本身即便只存在极小的偏见或倾向性，也会影响到数量庞大的个人。对于数据科学家而言，他们所面临的挑战就变成了发现、评估并化解可能使某些群体处于系统劣势的潜在偏见。

在保险行业，以下几个关于公平性的概念存在较强的相关性：

*精算公平*要求保险公司以相似的方式对待相似的风险，以便个人支付的保险费与实际发生的风险相对应。

*非歧视*指，个人支付的保险费不应由其无法影响的因素决定，尤其不应由那些社会保护群体独享的因素决定。非歧视与群体公平性等概念相关。那些具备保护属性（如性别、种族、性取向等）的群体，他们的目标是收到与那些不带保护属性群体类似的对待或结果，避免因某项敏感属性而受到区别对待。群体公平性弥补了差别影响的缺陷，差别影响即决策结果不恰当地伤害了或使某些具备特定敏感属性的人受益。

*互助或互惠*常被认为是保险的一项核心特征。在人工智能系统和数据分析技术的推动下，保险越来越注重个体化、个性化，这一趋势可能会剥夺特定群体应有的权利，比如向特定群体收取难以负担的保险费，或干脆拒绝提供保障。

不同公平性概念在不同险种和各国管辖范围间的相对重要性有所差异。例如，在许多国家，互助被认为是医疗保险的一项关键特征。

为了更好地发现、评估并化解潜在的偏见，保险公司需要在每次使用人工智能技术时根据不同情境确认公平的具体概念。在保险行业，公平性已不是一个新概念。所以应充分利用好现有的框架和规范（如精算伦理）。

<sup>1</sup>算法代理可以理解为某种近似模型，这些模型模拟算法的产出结果，因而更易理解。

## 建议

为了在保险公司内部推广运用负责任的人工智能系统，保险公司应考虑采取以下三项措施：

### 1) 建立使用人工智能技术的内部指导方针和政策

就运用人工智能技术确立内部指导方针和政策有助于在保险中使用人工智能时提升利弊权衡的认知。保险公司应制定和采取相应的指导方针和政策，包括处理透明性、可解释性和公平性问题的原则。尤其是，指导方针应有助于形成基于各个案例的异同阐明好处和风险的评估办法。精算师、风险管理人员、数据科学家和数据保护专员应紧密配合，共同研发、执行这些指导方针和政策。

在此过程中，保险公司可以采取一种以风险为基础的方法来管理使用人工智能技术，着重关注那些对个人有显著影响的使用情况。影响的显著性指的是决策对个人的影响结果，因人工智能具体的使用情况而有所不同。

### 2) 采用合适的治理框架

为了更负责地运用人工智能技术，保险公司应建立一套完善的治理框架，使每一级的决策者，包括个人、委员会和部门各司其责。决策者们应掌握必要的资质、技能和专业知识，而公司应设置有效的流程，包括触发流程和升级流程等。

治理模式种类繁多，各有利弊，宜参照公司自身架构及企业文化选择最合适的模型。

### 3) 制定并出台内部培训项目

最后，为了使全体员工更负责地运用人工智能技术，公司需要各职能部门、各层级的管理人员了解人工智能的好处和风险。为此，保险公司应考虑制定并推行全面的培训计划，阐明使用人工智能技术的好处和风险，以及各自内部的指导方针和政策。这些培训计划能更好地面向不同层级、不同职能的决策者，包括保险代理人和其他面向客户的员工。

