

数据使用的透明度和避免算法结果出现误差都非常重要

## 保险业的数据驱动：从数据错觉到数据解决方案 第二部分

这篇文章一共由两部分组成，旨在帮助那些希望利用数据和科技成为客户的伙伴而不仅仅作为理赔付款人的保险公司回顾一些重要的概念和考虑因素。当中包括道德考虑、准确理解数据的内容、数据滥用和合理使用智能数据的意义，以及整合不同来源数据的重要性，最终有助于我们能够掌握数据的真正潜力。

[点击这里](#)获取文章的第一部分或者继续阅读。

### 科技已经问世、随时可用，但也在不断发展

从保险业的角度来看，科技和数据的融合为重新定义我们的产品和客户体验提供了非常好的机会。

正如在第一部分中提到，数字健康设备被归类为物联网（IoT）设备。这表示它们配置了传感器、软件、网络连接以及收集和交换数据的能力。<sup>1</sup>

<sup>1</sup> See <https://internetofthingswiki.com/internet-of-things-definition/> Accessed 28 October 2019

<sup>2</sup> See Balasubramanian, R., et al., McKinsey & Company, New York, NY, USA. Insurance 2030 – The impact of AI on the future of insurance. , Apr.2018

麦肯锡（McKinsey）指出，在未来几年中，我们将会看到个人联网设备数量的大幅增长。“现有设备（例如汽车、健身监测仪、家庭助手、智能手机和智能手表）的普及率将继续快速增长，服装、眼镜、家用电器、医学设备和鞋子等的新兴类别也将加入增长行列。”<sup>2</sup>

他们还指出，“这些设备产生的大量新数据将使运营商能够更深入地了解自己的客户，从而衍生出新的产品类别、更个性化的定价和提供更加实时的服务”。

无论如何，站在汉诺威再保险的角度，我们认为人寿保险公司只有清楚地认识到自己需要解决哪些问题，才有可能针对性地收集到准确的数据。

### 数据需要使用相同的语言、或者可以被编译，并且必须是流动的

下面是摘自《财富》杂志在 2019 年 3 月发表的一篇文章，《点击 100 次后死亡：电子健康记录（EHR）哪里出了错》。<sup>3</sup>

<sup>3</sup> See Schulte F., Fry E., Death by 1,000 Clicks: Where Electronic Health Records Went Wrong, March 18 2019

“两天来，一位年轻律师的头一直疼痛不止，发烧使他头晕目眩，他需要很费劲地告诉急救接线员他的住址。

医院的一名医生怀疑他患了脑膜炎，并为他进行了脊椎穿刺。第二天，一名传染病专家在医院的电子健康记录（EHR）系统中录入了一个关键的实验室检查医嘱——检查脑脊液中是否存在病毒，包括单纯疱疹病毒。

尽管该医嘱显示在了 Epic 电子健康记录的屏幕上，但实际并未发送到实验室。根据患者罗尼斯基（Ronisky）在 2017 年 2 月向洛杉矶郡高等法院提起诉讼中披露的内容，Epic 软件并未彻底与实验室软件“对接”。患者声称，他的检查结果和诊断因此被耽搁了好几天。在此期间，由于疱疹性脑炎他遭受了不可逆转的脑损伤。”

随着数据源的激增，一个重要的考虑因素是我们正面临着形成数据孤岛的风险，有价值的无法在创建目的之外的领域中应用。例如，虽然血糖连续监测应用程序可以记录数据并被可视化后用于临床用途，但以保险等其他目的的数据访问还有待于开发。

上面提到的电子健康记录（EHR）例子突显出数据的重要性。设计 EHR 的初衷是在您允许的情况下，任何医疗保健提供者都可以访问到您的健康记录，因为这些记录都以电子格式存在。EHR 平台还可以用于申请血液检查，查阅不同的检查项目（例如 X 线等）。目的是提供高效的、经济的护理。虽然理论上这是一个显著的进步，但问题可能甚至已经出现。

除了平台本身非常复杂之外，EHR 还由不同的公司承建，使用它的医生们看起来只需要输入很少的信息。因此，除了由于制造商造成的差异之外，它自身也存在缺陷，并且提供的用户体验也较差。出现过的严重问题包括已下达的检测医嘱未通知到实验室、以及开出的处方药剂量或用药时间不正确。这应当引起你的注意，为客户制定一个非常清晰的数据采集和使用计划来改善流程非常重要，而不是使情况变得更糟糕。

让不同的平台和系统之间相互通讯和共享信息被称为互用性，具体是指“不同信息系统、设备和应用程序（系统）在机构、地区和国家的范围内以协调的方式访问、交换、集成和使用数据的能力，提供及时和无缝的信息传输，全局性地优化个人和族群的健康状况”。<sup>4</sup>

我们可以期待通过互用性缓解诸如上述提到的突出问题。物联网医疗设备也将很快为 EHR 提供数据，从而支持更加全面的护理。

这同样有可能成为人寿保险公司的机遇，但也反映出当今的医疗数据的产生已进入到数字化阶段。

人寿保险公司长期以来一直是医疗信息的使用者。在过去，医疗数据是在纸质文件上创建和存储的，因此我们按它被创建的格式查阅。在某种程度上，现在以及将来的医疗数据（无论类型或来源）已经在以数字化的方式创建和存储。我们作为一个相关行业需要按它被创建的格式进行访问。

展望未来的 10 到 15 年，大部分购买保险产品的人群、以及保单有效期更长的人群将会是千禧一代（大致对应 80 后、90 后人群）和 Z 一代（大致对应 95 后和 00 后人群）。对这些数字原住民来说，他们了解事物的唯一途径就是数码世界中的交流。这表示由于这些人群的互动方式数字化，因此他们对于数据访问的授权往往是根据公司提供有用产品的能力以及作为个人数据受托人所拥有的声誉来决定。

## 数据的目的

数据最终必须服务于特定的目的。在保险的领域中，这可以被广泛地描述为以最优惠的（匹配风险的）价格、以自动化的（尽可能低争议的）方式向需要的人群提供保障，即使他们可能没有察觉到。

为了实现这一目标，数据必须在诸如发挥人工智能优势的方式下加以利用。此前我没有特别提到人工智能，是因为首先需要认识数据本身。

<sup>4</sup> See HIMSS What is Interoperability, Health information Management Systems Society

根据如何与人寿保险公司的宗旨产生共鸣，以下罗列了一些关键的概念（没有单一定义）：

**人工智能 (Artificial Intelligence)**：“计算机科学的一个领域，致力于解决与人类智能有关的认知问题，例如学习、问题处理和模式识别”。<sup>5</sup>

**机器学习 (Machine Learning)**：“作为人工智能的一个分支，机器学习旨在通过数据、观察以及和世界的互动向计算机提供知识。所获得的知识可被计算机正确地应用到新的设置中”。机器学习可以受到、也可以不受到监督。<sup>6</sup>

**深度学习 (Deep Learning)**：深度学习是机器学习的一个分支，其中受人脑功能启发而开发的算法是类神经网络 (artificial neural networks)，它将大量的数据作为学习对象。和我们吸取经验的过程相似，深度学习算法会重复执行一个任务，每次都会稍作调整以改进结果。<sup>7</sup>

在以上的定义中，我只会提到人工智能 (AI)，但通过它的层次结构，也会涉及机器学习和深度学习。

相信许多人都听说过“垃圾输入、垃圾输出”这句话，对于人工智能来说尤其是如此。它并不是一个万能工厂，可以接收任何数据并输出智能、有意义和可被操作的结果。

事实上，即使是输入优质的数据也会出现误差、并且出现意料之外的结果。《科学》杂志最近的一篇文章中就曾提到，某家健康保险公司所使用的健康算法在相同的健康评分患者中得出了黑人患者比白人患者更为不利的结论。“产生这一偏差的原因是该算法是基于医疗保健费用而不是疾病进行预测，但实际上患者获得医疗服务的机会并不均等。花在黑人患者上的总费用相比白人患者更少。因此，尽管通过某些措施的预测校准，医疗保健费用可以有效地代表健康状况，但仍有可能出现较大的族群差别。”<sup>8</sup>

虽然这只是一个使用数据进行医疗决策的案例，但它体现出误差对结果的影响。如果算法的构建不充分，误差也将会很容易影响人寿保险业务中的风险评估算法。

尽管拥有良好的初衷，但我们还需要认识到注意力经济以及数据使用不当、或者使用不当数据可能对行业造成的破坏性影响。鉴于数据的类型和来源千差万别，从正确的来源中选取准确的个人数据、并将其转化为有用的产品回馈给数据提供者，这种能力对保险公司至关重要。根据通用数据保护条例 (GDPR) 的宗旨，我们也不应忽略实现自动化目标的过程需要准确和公平。

根据 GDPR 和类似数据法规的规定，投保人有权不接受自动决策的结果。与之相对应的、同时也是汉诺威再保险所秉持的是，我们不会支持无法被解释的自动决策。要使决策能够被解释的唯一方法，首先，是需要了解用于构建你的人工智能保险模型的数据；其次，还需要知晓你的人工智能模型对该数据进行了哪些处理。

## 总结

由于外部数据源的迅猛扩展，识别出真正具有保险用途的高质量数据并不容易。因此，能找准原因确定需要什么数据、去解决什么问题就显得更加重要。同样重要的，还包括明白获取数字化的健康数据使得纵向健康管理成为可能。这些信息将有可能解决无形性价值悖论 (VIP) 的问题。

正如麦肯锡 (McKinsey) 曾准确地指出，“未来人工智能创造价值，将需要运营商整合来自整个机构的技能、技术和洞察力，以提供独特和全面的客户体验”。对于保险公司而言，我们特有的行业知识和经验契合内部数据技能是我们的优势，但也必须意识到何时需要外部专家以确保开发出准确的、公平的和可被解释的人工智能解决方案。

<sup>5</sup> See Marr B., Forbes, February 14 2018, The Key Definitions of Artificial Intelligence That Explain Its Importance

<sup>6</sup> See Faggella D., The Rise of Neural Networks and Deep Learning in Our Everyday Lives – A Conversation with Yoshua Bengio, Emerj, February 19 2019

<sup>7</sup> See Marr B., Forbes, February 14 2018, What Is Deep Learning AI? A Simple Guide with 8 Practical Examples, October 1 2018

<sup>8</sup> See Obermeyer, Z., et al., S. Science, 366(6464), pp.447-453. Dissecting racial bias in an algorithm used to manage the health of populations. 2019

寻求能够解决关键问题的准确数据源，铭记获取和处理数据并将其转化有价商品回馈给投保人的复杂性，这些都将成为寿险公司的核心能力。

数据使用的透明度和避免算法结果出现误差都非常重要，不仅是由于数据保护和使用法规的约束，还因为作为值得信赖的财务安全提供商的保险公司，其声誉有赖于以合适的价格大众提供恰当的保障。

“变革的步伐从未如此之快，但也不会再如此之慢”

——贾斯汀·特鲁多，世界经济论坛 2018

最好我们所有人都赶上数字化转型的列车，一旦它离开车站就将难以追赶。

我们乐于为您的数字化转型提供建议。hr | equarium 计划 (www.equarium.com) 将保险技术公司和保险公司联系在一起，外加通过 hr | QUIRC 和 hr | Reflex 所呈现出的我们在自动化和外部数据集成方面的经验，我们已准备好与您一同前行。

## 联系人



**Dr Matthew Procter 医生**  
Medical Doctor  
电话+27 11 481 6729  
matthew.procter@hannover-re.co.za

在 LinkedIn 上关注我们，了解最新的寿险与健康险资讯。



## 参考资料

Balasubramanian, R., Libarikian, A. and McElhaney, D., McKinsey & Company, New York, NY, USA. Insurance 2030—The impact of AI on the future of insurance. , Apr.2018

Faggella D., The Rise of Neural Networks and Deep Learning in Our everyday Lives – A Conversation with Yoshua Bengio, Emerj, February 19 2019, Retrieved from <https://emerj.com/ai-podcast-interviews/the-rise-of-neural-networks-and-deep-learning-in-our-everyday-lives-a-conversation-with-yoshua-bengio/>

HIMSS, What is Interoperability, Health information Management Systems Society, Retrieved from <https://www.himss.org/library/interoperability-standards/what-is-interoperability>. Accessed 1 November 2019

Marr B., Forbes, February 14 2018, The Key Definitions of Artificial Intelligence That Explain Its Importance, Retrieved from <https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#1265c27d4f5d>.

Marr B., Forbes, February 14 2018, What Is Deep Learning AI? A Simple Guide with 8 Practical Examples, October 1 2018, Retrieved from <https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/#2dab8f778d4b>

Obermeyer, Z., Powers, B., Vogeli, C. and Mullainathan, S. Science, 366(6464), pp.447-453. Dissecting racial bias in an algorithm used to manage the health of populations. 2019

Schulte F., Fry E., Death by 1,000 Clicks: Where Electronic Health Records Went Wrong, March 18 2019 Retrieved from <https://khn.org/news/death-by-a-thousand-clicks/>.

Understanding Internet of Things, Retrieved from <https://internetofthingswiki.com/internet-of-things-definition/>. Accessed 28 October 2019

The information provided in this document does in no way whatsoever constitute legal, accounting, tax or other professional advice. While Hannover Rück SE has endeavoured to include in this document information it believes to be reliable, complete and up-to-date, the company does not make any representation or warranty, express or implied, as to the accuracy, completeness or updated status of such information. Therefore, in no case whatsoever will Hannover Rück SE and its affiliated companies or directors, officers or employees be liable to anyone for any decision made or action taken in conjunction with the information in this document or for any related damages.

© Hannover Rück SE. All rights reserved. Hannover Re is the registered service mark of Hannover Rück SE