

百分点：智能+智慧，打造动态知识图谱

本刊记者/朱琨

动态知识图谱是串联现实世界和数据世界的核心，在发展动态知识图谱的过程中，百分点提出机器智能+人脑智慧这一概念，并利用这一概念来解决动态知识图谱中海量的数据挖掘和复杂业务的识别等问题。

信

息技术发展的脚步已逐渐加快，单一的技术及静态的数据已无法满足业务的发展需求，所以人们更加注重融合技术和动态数据的处理和运用。因此，需要在动态数据中进行信息的挖掘和分析，将碎片化的数据进行串联，从中获取并转化为对企业有价值的知识，最终来提升企业的商业决策能力和业务效率。这一需求就促使了知识图谱这一融合技术向各领域的应用场景进行延伸。

人、物、组织、时空、虚拟标识，建立动态知识图谱

北京百分点信息科技有限公司（以下简称百分点），作为国内领先的大数据和人工智能技术和产品提供商，自2009年成立至今，一直专注于大数据操作系统和智能认知产品以及智能决策应用场景的搭建，走在行业的前沿。为了满足企业客户数字化转型的核心需求，在2014年，百分点推出了知识图谱产品，紧接着在2015年推出知识图谱引擎，成为国内最早将知识图谱应用于行业业务的企业之一，目前在媒体出版、公共安全、金融、电商等行业都构建了对应的知识图谱，构建出了行业最大的行业知识图谱库，实体数达到数千万的量级，关系数达到几十亿的量级。

据百分点高级研发总监黄伟介绍，在2015

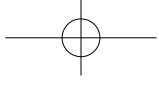
年开始百分点便深耕公共安全领域，并在落地实际业务场景中发现，对单一人物的刻画无法形成完善的公共安全解决方案，而更加注重人与人之间的关系。因此，百分点将人物标签系统进行了升级，增加了事件、关系等维度，从而形成行业知识图谱库。以此为基础，百分点研发创新了一项领先的知识抽取和知识融合技术，通过“机器智能+人脑智慧”，将现实世界中的“人、物、组织、时空、虚拟标识”映射到数字世界中，自动构建他们之间的关联关系，支撑用户展开分析和智能决策，这就是百分点的动态知识图谱技术。

“百分点动态知识图谱技术中很重要一点，就是当新的数据源进来，可以动态调整本体模型，使得知识图谱引擎保持正常、稳定的运转。”黄伟指出。

机器学习拓宽了人工智能的应用范围

百分点自成立之初就着力发展人工智能技术，采用了大量的机器学习技术来优化个性化推荐效果。黄伟认为：“打造个性化服务，重要的是通过分析个体行为从而理解并预测个体的需求。比如在电商领域，对商品信息及个体对商品的行为信息进行分析，以达到更好的推荐效果。”

同时，黄伟提出，若要采用机器学习技术对数据进行分析，应当满足两个先决条件。第



一，企业有明确的业务优化目标；第二，企业有支持优化的数据基础，这个数据基础除了高价值密度的结构化数据外，低价值密度但体量巨大的非结构化数据同样重要，包括文本、语音、图像、视频等。而传统的技术无法对非结构化数据进行有效的处理，所以百分点使用机器学习技术来训练模型，让机器自动对海量非结构化数据进行处理，最终使机器学习能够将非结构化数据转化为结构化数据，而这一做法也使百分点逐渐拓宽了人工智能的应用范围。

黄伟补充道：“在百分点的发展初期，就率先尝试利用机器学习以及NLP技术，提高推荐效果，随着业务技术的发展，NLP逐渐向知识图谱技术延伸，有了目前的动态知识图谱技术，使得百分点可以快速构建行业知识图谱，辅助行业智能决策。”

机器智能+人脑智慧优化模型

动态知识图谱是串联现实世界和数据世界的核心，在发展动态知识图谱的过程中，百分点提出机器智能+人脑智慧这一概念，并利用这一概念来解决动态知识图谱中海量的数据挖掘和复杂业务的识别等问题。但是百分点是如何解决机器智能+人脑智慧所涉及的机器与人脑互相协调问题的呢？

在公共安全领域，动态知识图谱的应用十分广泛。黄伟认为，虽然人工智能技术日渐强大和成熟，但是目前在人工智能水平还远没有到能够替代人工的状况下，我们提出机器智能+人工智慧相结合的方式，以机器+人脑提效为目标，更好地解决公共安全行业的业务问题。

目前，动态知识图谱的模型大致分为两类：预知模型和场景模型。那么，如何将这两类模型更好地融入到动态知识图谱的技术中，也成为了百分点所要考虑的问题。

黄伟介绍：“对于预知模型而言，它需要依靠行业业务的发展方向，并吸纳新型业务与技术进行搭建，所以预知模型需要在行业中不断进行积累才可以逐步完善。针对预知模型的这一特性，百分点将动态知识图谱所涉及的所有数据进行整合，通过不同的应用场景建立不同的模型，逐一添加到预知模型所涉及的应用场景中，同时将预知模型作用到BI平台上。”

客户的事就是自己的事

百分点历经十年的发展，俨然是中国大数据和人工智能企业的缩影。在近些年中，百分点依靠自身独特的技术及人工智能场景应用能力，成为了大数据和人工智能领域的独角兽。百分点将2B作为自身商业模式的主战场，但是无论是国内或是海外的业务拓展，百分点更多的体现了一个“蹚”字。百分点认为，若要打造2B市场，就要非常重视场景落地的技术和经验。

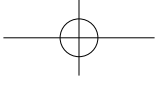
“所以，百分点在服务于任何一家传统企业时，都将客户的事当做自己的事来做，而不是一锤子买卖。”在这一点上，黄伟的体会尤其深刻，百分点一直致力于利用技术和产品来解决客户的实际问题。在近些年中，百分点在国内外市场双轮驱动下，大量拥抱商业化和场景化的落地模式，并依靠全栈的大数据+AI技术产品和场景化应用能力，服务了众多行业的上万家客户。

访谈实录

Q：中国大数据产业生态联盟、《软件和集成电路》杂志社

A：北京百分点信息科技有限公司高级研发总监黄伟

Q：动态知识图谱的价值体现在哪些方面？



A: 动态知识图谱是一个偏底层的技术, 我们在开发动态知识图谱时, 经历了非常漫长的迭代过程。所以我们在开发动态知识图谱的过程中, 对该项技术进行了权衡, 我们要考虑的是, 我们开发的动态知识图谱技术能为客户带来哪些价值?

第一, 我们所提倡的是灵活性, 客户积累的任何数据, 都能进行对接。可动态修改数据拉通规则和融合配置, 整个过程是实时的, 无需重启系统, 无需重新导入数据。

第二, 实现对任意不同领域进行建模, 对同一领域可应用多种不同的建模方式, 数据模型可动态扩展和调整, 最终保证该项技术可以快速部署到任何场景中, 可到现场帮助客户直接安装。

第三, 基于动态知识图谱和行业业务模型, 具备自优化和自适应能力, 实现知识进化, 优化更新知识表示, 支持复杂业务问题的自动识别。构建面向任何问题域的知识图谱, 帮助企业实现智能决策。

Q: 百分点动态知识图谱的核心是什么?

A: 动态知识图谱的核心是: 知识图谱+动态本体。知识图谱实现数据到知识的升华。运用知识图谱将数据聚合到一起, 大幅提升单数据源能够发挥的价值。动态本体实现新增数据源的问题。接入新的数据源不影响已有图谱的线上使用, 更改融合规则数据上实时体现。

Q: 动态数据融合的规则有哪些? 如何将规则变成现实?

A: 数据融合的规则都是业务人员在产品中设置的规则, 当不同源的信息发生冲突时, 需要规则中指定信息权重来动态改变融合结果。融合规则的改变对数据融合结果是实时体现的。

Q: 百分点的动态知识图谱可以自动构建实体与时空映射的关系, 如何做到“动态”?

A: 动态知识图谱的“动态”有两层含义。第一, 动态本体。通俗来说, 数据的语言其实就是方案, 我们在打造本体方案的过程中, 必须更加了解人认识现实世界的方式, 这是逻辑概念。第二, 动态实例, 真实的世界是时刻变化的, 反映真实世界的图谱亦是动态变化的, 所有对象具有很强时间属性。

Q: 如何解决海量数据的存储与分析的难题?

A: 处理海量数据, 需要运用大数据分布式存储和计算框架来做底层支撑。从对内容分析的角度上看有两大类难题。

第一部分: 对结构化数据的处理。在安全行业中结构化数据有很多种, 这些结构化数据一般是高价值、高密度的数据。

如何快速地把不同种类的结构化数据进行融合, 并且能支持未来新增融合数据源, 这确实给我们带来了许多挑战, 在找寻到解决方法后, 也整合到了动态知识图谱的底层技术中去。

第二部分: 对非结构化数据的处理。互联网上存在着大量的信息, 绝大部分都是非结构化数据。如何把非结构化数据进行整理, 最终使非结构化数据产生更多的知识, 并找到细微的知识聚合在一起作用到业务中, 这也是我们一直在重点解决的问题。

行业应用案例

案例名称: AI技术助力山东某市构建新型治安防控网络

核心特点: 百分点DeepFinder高效整合集成了当地近百类数据资源、几十亿条数据, 并依

图 “数据双胞胎” 的核心：动态知识图谱



托“动态知识图谱”技术，对每天新增千万条数据进行动态融合，实现对公安全要素数据高效整合和治理，通过对实时数据流计算，主动发现、识别个人和群体行为的异常，并根据风险预警模型分级预警，指导辖区警务部门及时调整警力资源部署，将风险隐患消除化解在萌芽状态，有效降低警力人工支出，解放警力资源，提升公安工作大数据应用与分析研判能力，实现智能决策。

应用解读：

随着基础设施建设不断完善和各类数据不断增长，从海量多源异构的数据中进行价值信息的获取，还停留在利用经验积累进行人工分析的阶段，庞大的数据量对公安机关案件破获带来困难。对各类风险隐患的敏锐感知、精确预警，可有效遏制犯罪行为的发生，提升人民群众安全感。

上台峰会期间，百分点为峰会重点枢纽区域的警务部门，提供大数据挖掘分析技术和智能安全分析平台。实现对可疑人员、可疑车辆、风险隐患的智能识别、动态轨迹追踪和精确预

警。构建新型社会治安防控网络，全面感知公共安全态势，实现对各类潜在风险隐患的主动预测预警预防，防患于未然。

应用价值：

1.AI技术构建新型社会治安防控网络

帮助公安机关运用大数据技术对潜在风险隐患智能识别、动态轨迹追踪和准确预警，构建新型社会治安防控网络，全面感知公共安全态势，实现对各类风险隐患的敏锐感知、精确预警。帮助公安机关实现以数据为关键要素的数字侦查。

2.推进国家治理体系和治理能力现代化

通过对公安全要素数据高效整合和深度治理，不断增强对各类治安要素的掌控制，运用数据分析模型，关联发掘重大热点难点问题，提高对各类社会矛盾的发现预警能力，及时排除、预警各类矛盾风险，推动从依靠直觉与经验决策向依靠大数据决策转变，从而维护群众的生命财产和公共安全，形成有效的社会治理、良好的社会秩序，帮助提升社会治理和管理水平，推进国家治理体系和治理能力