

MATRIX 2.0 绿皮书

1. 背景

数据:未来社会最重要的资产

随着互联网与移动互联网的高速发展,我们的世界已经逐步步入数字经济时代,这是社会的一次重大变革,然而随着区块链技术的出现,这次变革将上升到一个前所未有的新高度。

对于任何一个社会形态,都满足以下三个原则:

1. 资产是指任何公司、机构和个人拥有的任何具有商业或交换价值的东西:

无论是黄金这样的传统资产,还是我们购买的类似于理发这样的服务,都是资产,在数字经济时代,尤其是随着人工智能技术的发展,数据的商业及交换价值越来越高,成为了一种全新的以及在未来最重要的资产;

2. 资产价值的基础是确权:

无论是纸币、贵金属,还是数据或服务,资产的价值都基于其权益被确认的基础之上,无法确权的资产无法具备价值或者其价值将大幅缩水;

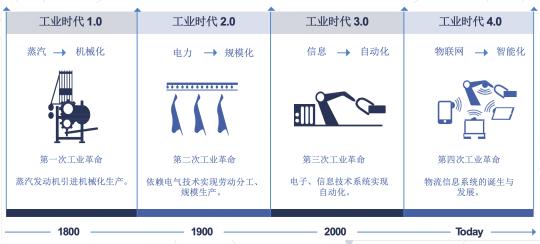
3. 社会经济发展的动力是资产价值的增长:

人类历史、科技和生产力的发展,本质上都是在追求资产价值的增长。在原始社会,石头是原始资产的一种,我们把石头打造成武器,石头在工具化的同时,其资产的价值也实现了增长,一把石质武器的价值远高于石头本身;到了封建社会,人们开始从矿石中提炼金属,再打造成各种工具,资产的价值有了大幅的提升;而工业革命直接将金属变成了内燃机,资产的价

值有了飞跃性的增长,由此可见,社会发展的本质,是追求资产价值的增长。

在数字经济时代,数据是这个时代最基本的资产,我们通过人工智能技术,通过数据训练出AI模型,就像我们把石头加工成武器一样,AI模型直接将会在各个产业以及环境中应用,大幅提升人的工作效率、工业的生产效率甚至取代人的工作,这就是第四次工业革命。

生活在互联网时代,我们每天都会产生大量的数据,而基于这些数据,未来可能将产生更多高质量的模型,推动人工智能技术的高速发展,可以预见的是,在数字经济时代,谁的数据更多,质量更高,谁就能训练出更优质的模型,提供更好的服务,带来更多的用户和数据,所以数据资产的数量和质量,将取代传统资产以及移动互联网初期的流量,成为"数字经济体"的核心竞争力。



人类通过对资产价值增长的追求,推动全球四次工业革命

2. MATRIX 2.0 概述

我们运用AI技术解决了区块链的四个核心问题:速度慢、不安全、不易用以及资源浪费,从而打造了一个高性能、可靠的开源区块链平台,也就是说,在MATRIX 1.0时代,我们使用AI技术赋能区块链,打造了一个更好的区块链平台。

在未来,我们将以我们创造的AI赋能的区块链平台为基础,使用分布式的方式管理人工智能,打造一个更加有效,可以自我成长的人工智能生态体系。

如我们前面提到的,人工智能技术的三个基石分别是数据、算力和算法,我们将围绕着三个基石来打造一个基于区块链的人工智能经济体。



MATRIX 2.0生态体系

1) 数据权利受区块链保护

大数据已经是这个时代高速发展和取得显著成绩的一个产业,也是人工智能产业的前身,现在已经有很多成功的大数据公司,例如谷歌、例如亚马逊,但对于这些数据巨头来说,他们的数据仍然具有相当的局限性,我们把他们的数据称为"单维大数据",也就是数据的数量很庞大,但数据类型单一,例如谷歌只有用户使用谷歌产品的数据,亚马逊只有用户使用亚马逊产品的数据,同时,数据拷贝的门槛极低以及缺乏有效的保护数据隐私与权益的手段,数据巨头之间无法分享数据,这就创造出很多个"数据孤岛",大大地局限了人工智能技术的有效发展。

从另一个角度来说,每一个个体用户拥有丰富的数据,比如Owen,他拥有他使用谷歌的数据,也拥有使用亚马逊的数据,他拥有自己完整的通话记录和银行账单,他拥有自己每天的地理位置信息以及几十个APP的使用习惯,如此丰富的数据无疑是训练AI模型的高质量选择,但遗憾的是,Owen只拥有他自己一个人的数据,这个数据规模无法支持任何最小样本要求的机器学习,我们把这种数据称为"多维小数据"。

由上可见,真正能支持AI技术爆炸性发展的是"多维大数据",可如何能让每个终端用户都分享自己的数据?如何让这些数据巨头之间打破"数据孤岛"的桎梏,从而分享数据推动AI技术的发展?答案是区块链以及分布式存储技术。

我们将与合作伙伴在MATRIX的基础上,创造一个区块链数据平台,任何将数据分享到平台上的用户,无论是企业还是个人,都将获得奖励;同时更重要的是,区块链将对所有上链的数据确权,保证数据的唯一所有权;每份数据也将通过分布式存储以及加密技术存在链上,而避免直接被人拷走的风险。

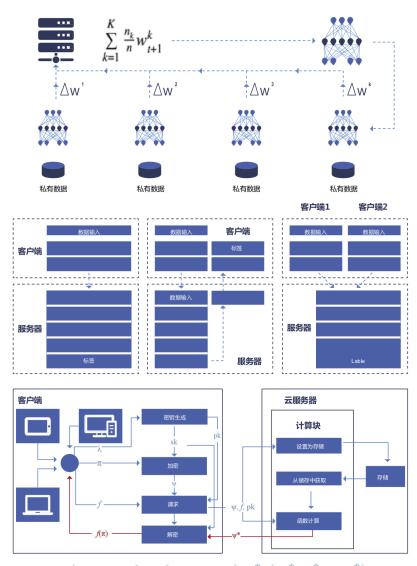
当然,仅仅是上传数据并不能完全挖掘数据的价值,工业4.0的核心就是将数据的价值挖掘到极致,所以在未来,任何使用

用户数据训练模型的行为,都需要向用户付费,同时这个模型在未来商业化运营中的收益,用户都将获得一部分分成。

如何在用户数据被使用的过程中仍然可以保证用户数据的隐私安全,我们将尝试以下两种核心技术:

联邦训练:

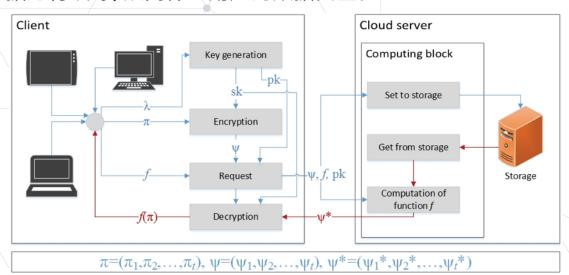
这是一个基于分布式存储的机器学习训练方式,每个终端将基于一部分数据分别执行训练任务,最后再将训练结果进行整合。在这种技术下,任何一个训练终端都无法获取用户的完整数据,从而保证用户的数据安全。



 $\pi = (\pi_1, \pi_2, \dots, \pi_t), \ \psi = (\psi_1, \psi_2, \dots, \psi_t), \ \psi^* = (\psi_1^*, \psi_2^*, \dots, \psi_t^*)$

同态加密:

同态加密是基于数学难题的计算复杂性理论的密码学技术。 对经过同态加密的数据进行处理得到一个输出,将这一输出 进行解密,其结果与用同一方法处理未加密的原始数据得到 的输出结果是一样的。即我们将不让数据科学家使用明文数 据进行训练,从而保证用户的数据安全。



同态加密示意图

2) 由区块链组织的云、雾和终端设备提供的计算能力

基于我们的设计,同时为了更加有效的提升整个网络的运营效率,以及终端用户的切实需求,我们打造了云-雾-终端计算架构,我们将整个算力层分为三层:

云计算

即MATRIX分布式AI算力分发网络,主要解决算力密集、反应时间较低的计算任务;

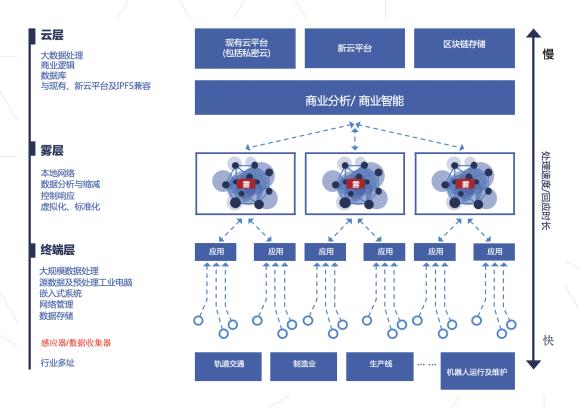
雾计算

用户本地网络的核心节点,主要解决算力要求相对较低、实时

性较强的计算任务;

终端计算

服务于特定应用的智能终端。



3) 使用区块链跟踪AI模型

AI的算法与模型是人工智能技术与我们最近的一层,因为所有的应用都需要基于算法与模型才可以实现。换句话说,算法和模型让海量的算力真正产生了价值,让冷冰冰的数据,具有了灵魂。

我们创新地使用区块链技术来管理和建立模型市场,主要为了达到三个主要目标:

1. 模型确权

和其他的计算机软件与服务一样,AI模型被模仿或抄袭的门槛很低,缺乏有效的知识产权保护手段是众多数据科学家长期挥之不去的梦魇,而如果大家都封闭自己的研究成果,又会对AI技术的发展产生阻碍。区块链技术可以完美地解决这个问题,我们使用区块链技术对科学家的模型进行确权,来保护科学家的知识产权。

2. 分布式训练

多样性是生物进化中非常重要的一个因素,对于人工智能也一样,基于区块链的去中心化的体系,更多的模型将被创造以及训练出来,我们相信去中心化的网络将像生物大爆炸一样,引导人工智能技术的爆炸性发展。

3. 可追溯的模型市场

区块链的另一个重要应用是溯源,我们基于这个特性打造了一个可追溯的模型市场,每位科学家创造出来的模型被用户或者应用开发者的每一次使用都将记录在区块链之上,我们用这个无法篡改的账本来记录科学家应该获得的每一笔收益,同时模型的真实使用状况的记录,也有助于科学家得到准确的反馈,从而在未来对模型进行更好的调优。

在保护数据科学家的知识产权的基础之上,我们提供了丰富的接口供直接用户或者应用开发者来轻松地使用或二次开发,大大降低AI技术为普通用户和各个产业服务的门槛。

3. MATRIX 2.0 架构

1) MATRIX 1.0

在AI技术的支持之下,我们打造了MATRIX 1.0版本,一个AI赋能的高效安全的去中心化开源公开平台(详情请见MATRIX白皮书1.0版本)。

MATRIX具有四个核心技术特点:

- 1. 更快速:使用聚类算法以及HPoW的共识机制在确保整个网络仍然满足去中心化特性的情况下,让MATRIX主网具有了更高的性能;
- 2. 更易用:使用AI技术创造了智慧合约(Intelligent Contract),在当前阶段,用户可以通过语音或文字输入生成特定类型的合约,在未来,我们希望通过机器学习积累,让用户可以通过语音即可生成自己想要的自定义类型的合约,让每一个人无论会不会编程,都可以使用智慧合约,使用区块链的核心应用;
- 3. 更安全:引入形式验证技术,让用户的合约和代码具有了更高的安全性,从而保护用户的合约和数字资产;
- 4. 更环保:使用AI计算来取代比特币的哈希计算,让PoW输出的结果真正具有社会价值,同时HPoW的共识机制下,大部分没有被选中的M节点的算力将可输出做有益的计算,例如癌症诊断、图形图像识别、金融建模等,从根本上解决了传统区块链挖矿浪费算力和资源的问题。

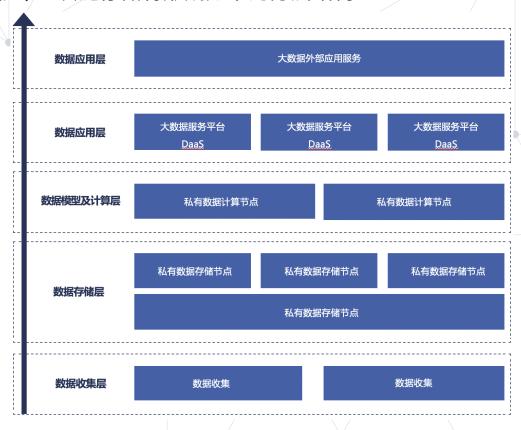
我们将基于MATRIX 1.0来打造MATRIX 2.0的生态体系,正是这样的一个高性能、安全、易用的区块链网络,让MATRIX 2.0

的实现成为了可能。

2) 区块链存储和计算节点

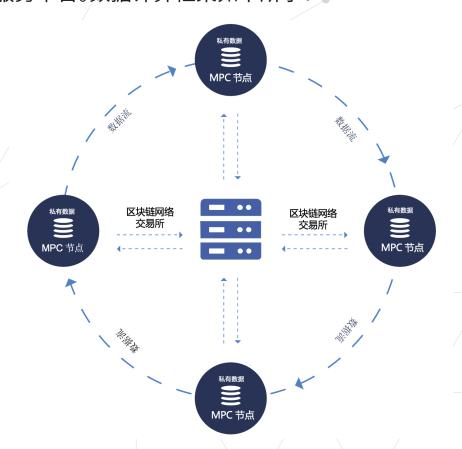
数据链的核心在于隐私保护和隐私存储,传统的区块链被设计为面向公众的公开数据存储,而对于半公开与隐私存储,则存在天然短板。当前时代是一个基于大数据与人工智能处理的时代,往往需要各个独立实体进行数据协作和共享,因此区块链未来的一个核心目标必然是保障隐私数据的安全,并支持各种形式的数据授权,保障数据的使用权和所有权分离。

通过设计支持数据安全和数据授权的区块链平台,最终能够实现:(1)保障用户数据隐私的前提下,实现大数据交易,并提高数据使用效率;(2)赋能行业大数据应用,支持各类企业数据协作与共享,实现企业人工智能的真实落地;(3)为大数据行业提供可持续发展的路径,实现数据产生、数据建模、数据应用和数据交易的分离与协作,避免传统意义上的数据孤岛。因此,区块链存储将形成如下的分层结构:



为解决隐私计算问题,区块链系统必须引入安全多方计算(Secure Multi-Party Computation,SMC)框架,该框架的核心是解决一组互不信任的参与方之间保护隐私的协同计算问题。SMC要确保输入的独立性、计算的正确性,同时不泄露各输入值给参与计算的其他成员。SMC通常应用于选举、投票、拍卖、秘密共享、门限签名等场景。

本方案将安全多方计算,作为区块链数据加密和验证机制的组成部分,区块链技术作为云计算、AI等基础设施平台的组成部分,结合零知识证明和其他密码学技术,构成了下一代通用计算服务平台。数据计算框架如下所示:



参与隐私计算的可以是任意一个区块链网络节点。该节点可以发起协同计算任务,也可以选择参与其他方发起的计算任务。路由寻址和计算逻辑传输由枢纽节点控制,寻找相关数据

同时传输计算逻辑。各个MPC节点根据计算逻辑,在本地数据库完成数据提取、计算,并将输出计算结果路由到指定节点,从而多方节点完成协同计算任务,输出唯一性结果。整个过程各方数据全部在本地,并不提供给其他节点,在保证数据隐私的情况下,将计算结果反馈到整个计算任务系统,从而各方得到正确的数据反馈。

3) 区块链操作系统

传统上,云计算设备分布于不同地理位置,通过统一的网管与分发平台实现,因此传统云计算实质上是与中心化紧密关联的。而新一代的区块链,除了完成分布式账本功能与隐私数据功能外,还需要经过各种功能整合,对外呈现出一个方便、灵活配置、即开即用、行为统一的服务方案,对此我们称之为区块链操作系统。

区块链操作系统的核心包含五个部分:(1)任务调度与算力分发;(2)算力资源管理;(3)支持隐私的云存储;(4)支持安全与路由控制的网络;(5)支持用户权限与分布式管理的安全机制。

对于开发者而言,区块链操作系统将提供智能合约虚拟机与oracle服务、多链、中间层、合约组建、Dapp组建等服务,并提供各种适合开发者使用的插件服务;对于使用者而言,区块链操作系统则需要提供额外的用户使用界面、智能化的人机交互系统和区块链应用交互系统。

针对区块链本身的特色,操作系统还需要提供一个定价系统,方便用户根据自身需求进行组件选择与服务选择。

4) 区块链支付模块

传统的数字货币支付,得益于底层区块链技术具有的溯源特性,可以轻而易举的知道每一笔钱的去向,数字货币的"留痕"和"可追踪性"能够提升经济交易活动的便利度和透明度。

同时,基于区块链技术能实现点对点的价值转移,通过资产数字化和重构金融基础设施架构,大幅度提升金融资产交易后清算、结算流程效率和降低成本的目标,并可在很大程度上解决支付方面面临的现存问题。数字货币支付可以降低跨境交易的成本,增加资本流动速度,这在一定程度上可以促进全球跨境贸易的发展。

但是区块链支付系统相对于传统支付的核心竞争力是用户数据,即用户账户数据、用户账本数据、信用记录。新的区块链支付系统核心是在用户数据挖掘与隐私保护。通过用户数据挖掘获得更深层次的数据,增强相对传统银行以及第三方的支付竞争力,通过隐私保护和脱敏技术,解除用户对数据使用的担忧,从而实现更好的隐私与公开的平衡。

4. MATRIX 2.0 生态

MATRIX 2.0的经济体系

与其他的公有链不同的是,MATRIX 2.0将不仅仅是一个公有链,除了未来链上的诸多Dapp带来的流通之外,人工智能技术对于链上算力和数据的需求以及相应的成本支出也将成为支撑这个生态价值的重要基石。在这个生态中,矿工为整个生态提供了算力设施,成为整个网络的硬件基础;带宽和存储的提供者为数据层及模型层提供了分布式的存储空间;数据提供者通过上传数据来打造多维大数据平台,这将成为整个平台的软件基础;而数据科学家在可确权的分布式网络下训练模型,将算力和数据转换成直接可为用户服务的工具。我们使用区块链技术和分布式存储技术来管理算力、数据以及算法,让MATRIX 2.0成为一个有无限进化可能的去中心化的"AI生命体"。



MATRIX 2.0与工业4.0

人工智能各产业中唯一作为单独概念被提出的就是"工业4.0",除了智能制造以外,AI智能维护及安全预警也将是工业4.0非常重要的组成部分,预计到2025年,AI智能维护及安全预警技术每年将至少为制造业节省6300亿美金。基于MATRIX团队的技术特长以及行业经验,预测性维护将成为MATRIX2.0前期重点拓展的一个行业方向。

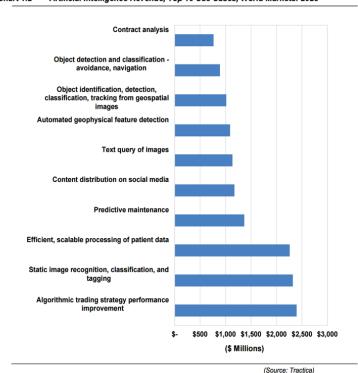


Chart 1.2 Artificial Intelligence Revenue, Top 10 Use Cases, World Markets: 2025

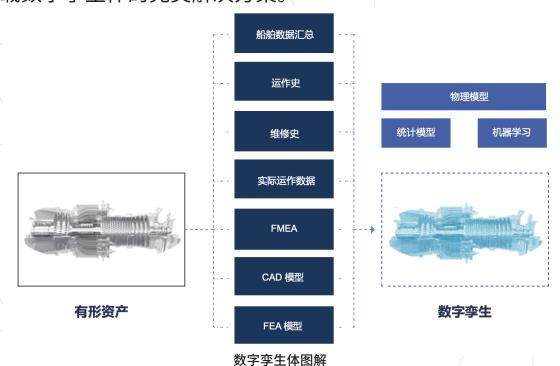
1. 数字孪生技术

智能制造时代亟需虚拟空间与物理空间的交互融合,业界需要一种既能承载产品设计信息,又能忠实反映真实物理产品的数字化定义方式。

数字孪生体是指产品物理实体的工作状态在信息空间的全要素重建及数字化映射,是集成的多物理、多尺度、超写实、动态

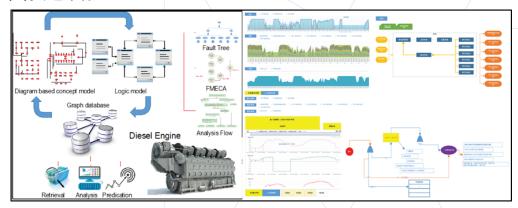
概率仿真模型,可用来模拟、监控、诊断、预测、控制产品物理实体在现实环境中的形成过程、状态和行为。

MATRIX 2.0中的分布式数据平台以及算法训练平台将成为承载数字孪生体的完美解决方案。



2. 故障诊断

根据数据异常或故障报警进行故障诊断;分析故障隔离、失效 模式和原因分析、波及范围;根据历史维修记录和作业指南形 成维修建议。

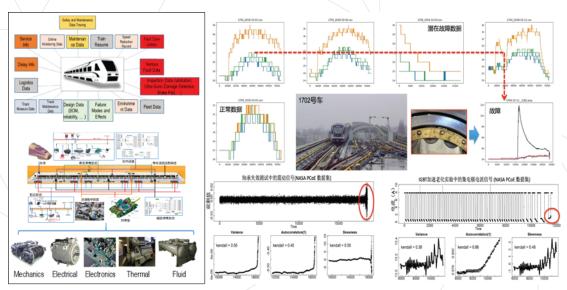


AI故障诊断推理系统



3. 故障预警

结合物联网技术来建设故障预警系统。整个系统分为三层,第一层数据收集与建模,是通过各种传感器收集设备部件的数据和信息,然后根据机器部件的基本物理建模来预测其何时开始磨损;第二层故障提示,基于收集到的数据盒模型,通过深度学习训练,可以提示机器发生的一般问题;第三层故障关联与预测,可将具体的数据标记、分类,然后通过深度学习将特定数据与具体故障联系起来。



已应用在中国轨道交通中的故障预警系统

4. 生产流程优化

首先,我们将设计目标以及各种参数(如材料、制造方法、成本限制等)输入到生成设计软件中;然后,AI将探索解决方案的所有可能的排列,并快速生成设计备选方案;最后,它利用机器学习来测试和学习每次迭代哪些有效,哪些无效,从而提升生产效率。

MATRIX 2.0与个人金融时代

随着社会经济的高度发展,人民可支配的财富日益增长,由于社会经济发展的动力是资产价值的增长,所以对于整个社会而言,资产流通速度的提升将大幅提升整个社会资产的价值,同时提升整个社会的运行效率,所以P2P的金融类创新层出不穷。但中心化的模式无法真正推动整个社会进入个人金融时代,因为一些核心问题无法得到真正的解决,如个人信息隐私问题、中心化中介机构的诚信问题,我们希望基于MATRIX 2.0的体系,真正推开个人金融时代的大门。

1. 去中心化个人征信隐私保护平台:

采用分布式存储的方式保存用户的个人征信数据,让个人征信数据不被侵犯,同时也解决了平台对用户征信查询困难的问题;

2. 智能征信评级系统:

采用最先进的小样本与非监督学习技术,可以有效准确地判断个人、小微企业等信息缺失标的的征信与风险评级情况,有效控制资金风险,同时有效扩展整个社会的投资标的与可使用资金规模;

3. 合约式分段管理资金:

资金将不会再放在中介机构的银行账户或托管账户,同时也不会一次性拨付给贷款方,整个资金都将通过链上的智慧合约保存,一整套的合约安全防护体系,保证合约的安全性,同时合约的条件式拨付也将大幅降低资金的使用安全风险。

MATRIX 2.0与智能医疗

预计到2025年,市场规模第三大的行业就是医疗行业,目前, MATRIX团队已经在医疗方面取得了卓越的成绩:MATRIX所 研发的AI癌症诊疗系统已经在多家医院开始应用,同时世界 首例小细胞肺癌诊断以及寿命预测系统的研发也已经接近尾声。

106 STARTUPS TRANSFORMING HEALTHCARE WITH AI



正如CBInsights统计中所显示,已经有大量的人工智能公司 迈入了智能医疗的领域,人工智能医疗诊断也是目前人工智 能科技最炙手可热的领域,但是,目前智能医疗离我们的生活 还有相当的距离,比如,没有足够多的优质数据、病人隐私保 护等问题阻碍了这个产业的高速发展。与其他的专门服务于 医院以及机构的AI公司或组织不同,我们希望基于MATRIX 2.0打造一个直接面向病人的人人可用的智能医疗平台。

1. 人人可用的智能医疗平台:

在这里,病人并不需要去分担巨额的医疗仪器的成本费用,还有AI计算的高额费用,病人只需要根据格式要求,在基于MATRIX主链的软件客户端上上传符合格式标准的病历数据或CT图像等,并使用MAN支付少量费用给提供算力的矿工,

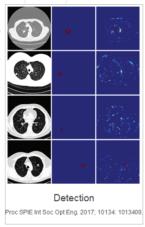
MATRIX网络将会基于我们的成果给病人相应的诊断结果,对那些医疗水平落后地区的患者来说,这大大降低了智能医疗的使用门槛;

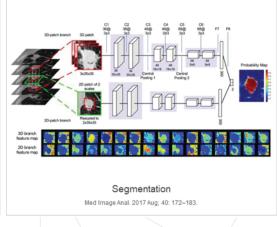
2. 真正安全的在线医疗平台:

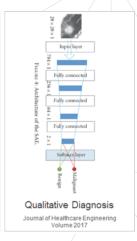
作为MATRIX 2.0基石之一的分布式存储体系,将在医疗产业 发挥巨大的作用,病人的个人信息上传后将会采用分布式的 方式存储在链上,从而避免被任何一个人窃取,同时分布式的 算力分析,也避免了病人的信息在诊断过程中被算力提供者 截取;

3. 可以不断成长的在线医疗平台:

随着大量真实病人的资料上传及就诊,MARTXI的医疗诊断平台将会基于这些高质量数据不断进化,从而不断提升诊断的准确度及可诊断疾病的类型;

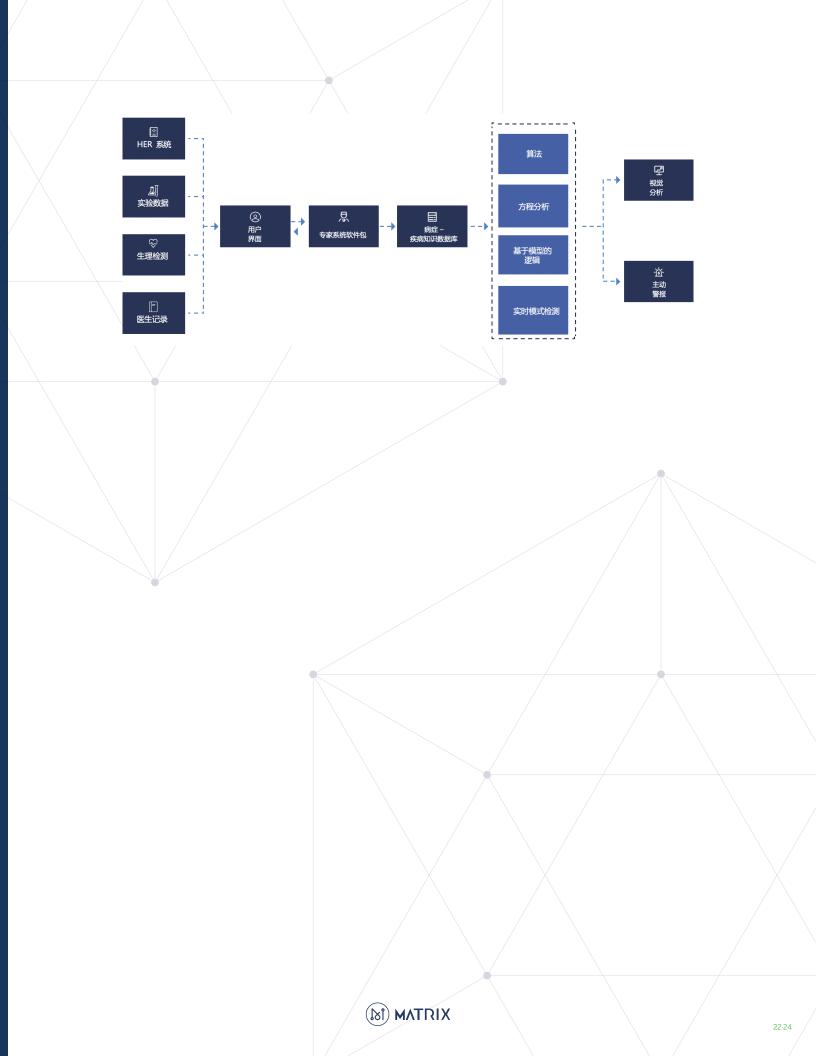






4. 高拓展性的智能医疗平台:

MATRIX 2.0将提供丰富的工具和基础模型,以及全世界最优秀的数据科学家的技术支持,允许优秀的医师直接优化数据和训练自己的模型,大大提升了平台的可拓展性。



5. 总结

我们正在目睹一个伟大的时代,大数据、人工智能和区块链技术正以前所未有的方式改变着我们的生活。机器学习应用程序提供了从数据中提取价值和创造社会利益的绝佳机会;我们还看到对保护个人数据和相关数字资产隐私的需求不断增加。MATRIX 2.0平台以先进的区块链、分布式计算和机器学习技术为基础,不断挖掘数据的价值、为数据安全和隐私保护奠定坚实的基础。持续探索人工智能在各领域的新可能,为区块链、人工智能技术的不断发展提供坚实的驱动力。

