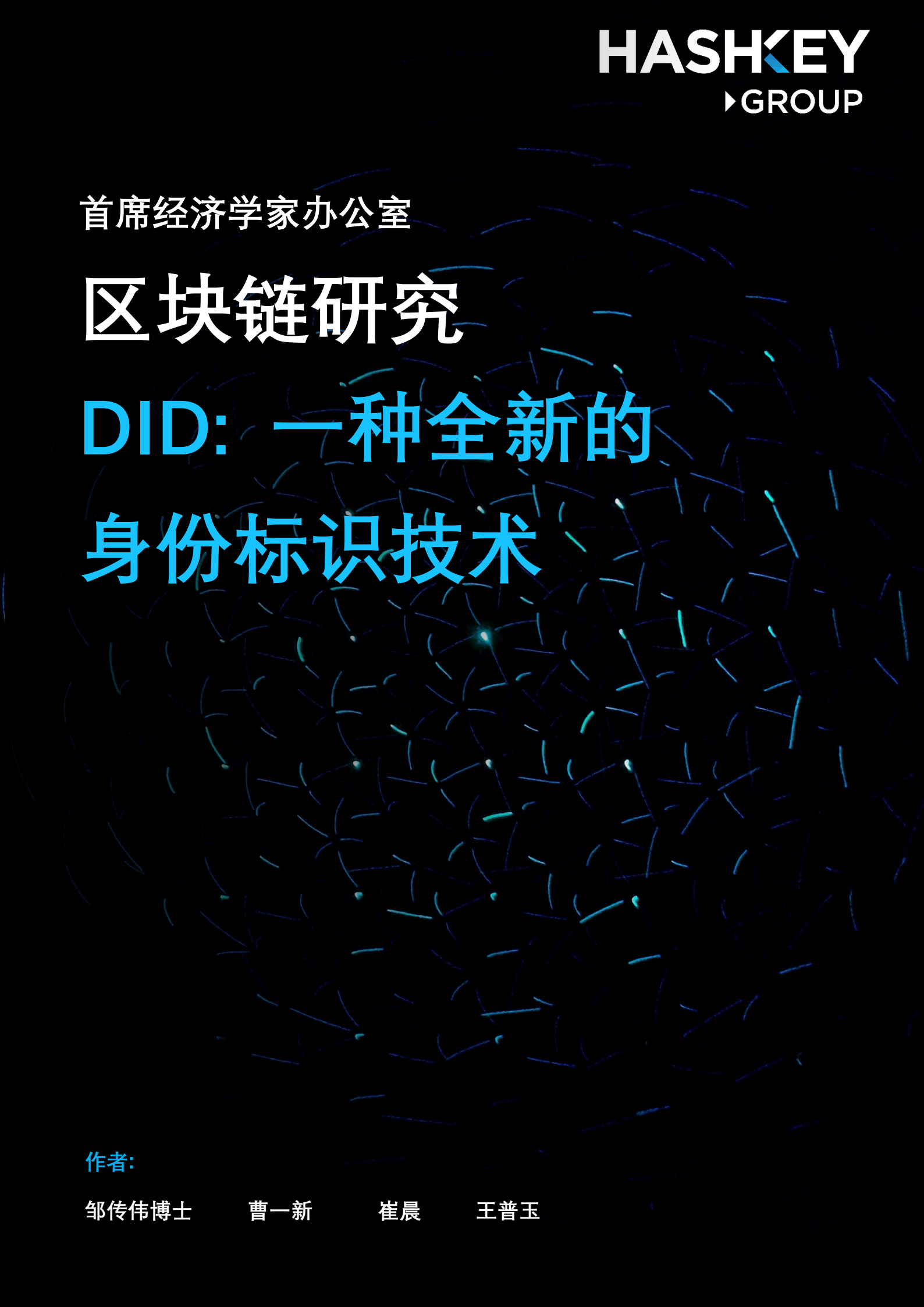
HASHKEY GROUP

首席经济学家办公室

区块链研究

DID：一种全新的身份标识技术



作者：

邹传伟博士　曹一新　崔晨　王普玉

摘要

近年来，大家对个人数据的隐私安全问题关注度明显上升，本文将从身份管理角度讨论该问题，主要包括四方面 :第一，传统身份管理方案有什么问题？第二，分布式数字身份（DID）是什么？技术实现过程是怎么样的？第三，DID 技术发展有什么瓶颈？第四，区块链技术和 DID 技术的结合能够擦除什么样的火花？

关键词：DID

第一章 身份管理

身份系统包括三个要素：身份、身份证明和身份验证。我们结合三个要素，从物理世界和互联网世界两个维度分别讨论身份管理。

1、身份系统

（1）身份

在物理世界中，每个人从出生就拥有独一无二的身份特征，包括外貌、体重、年龄、肤色、指纹等等；为了快速描述任何个体的身份特征，我们使用姓名作为代号，可以帮助大家快速识别他人并联想到关于对方的一切信息，这些内容被统称为身份。

与物理世界相对应的还有一个互联网世界，身份的概念完全不同。在互联网世界，用户完全可以根据自己的喜好设置想象中的“身份”，包括姓名、性别、身高、体重等；甚至可以随时更改这些“身份特征”，确切来说，此时的身份不同于传统意义上的身份，因为不具有唯一性和确定性。

（2）身份证明

在物理世界中，由人构成的系统变得日益庞大，为了便于中心化机构的管理，出现了身份证明。中心化机构根据不同人的身份特征签发了唯一身份证明，用于证明主体拥有某项资产的所有权或申明其享有某种社会权益，同时在不同个体及组织之间交互时，可以用于定责、纠纷追溯和信任保障。身份证明使身份的特征从隐性变为可视以及可追踪，例如政府签发的身份证、护照等，证明主体属于某个国家的身份以及享有某种权益；再比如驾驶证，能够证明某个身份具有车辆驾驶技能。

在互联网世界中的主体身份证明完全不同于物理世界。在物理世界中，身份证明与身份有着直接的关联，即通过身份证明就能映射到主体本人；但在最初的互联网世界中，身份证明和身份之间并不存在映射关系，不同主体只需根据设想中的身份特征（年龄、身高、姓名等）提交身份证明申请，而无需与物理世界身份特征保持一致，因此仅凭互联网身份证明是无法映射到主体本人。随着互联网世界的发展，匿名性和不可追溯性逐渐影响到了物理世界的治理和安全，多项规定要求平台方需做好用户实名制验证工作，这样就出现了互联网身份证明与物理世界身份证明映射的关系，进一步出现与主体身份映射的关系。

有意思的是，互联网用户的身份证明，需要依赖于物理世界身份证明的映射，来确定身份的唯一性和确定性。但互联网世界中的网站却完全不同，它从开始就有一套完整的身份证明体系，如图 1 所示的统一资源标识符 URI。每个网站拥有独一无二的域名，域名的签发（身份证明）是由国际域名管理中心统一管理，我国是由中国互联网络信息中心管理。

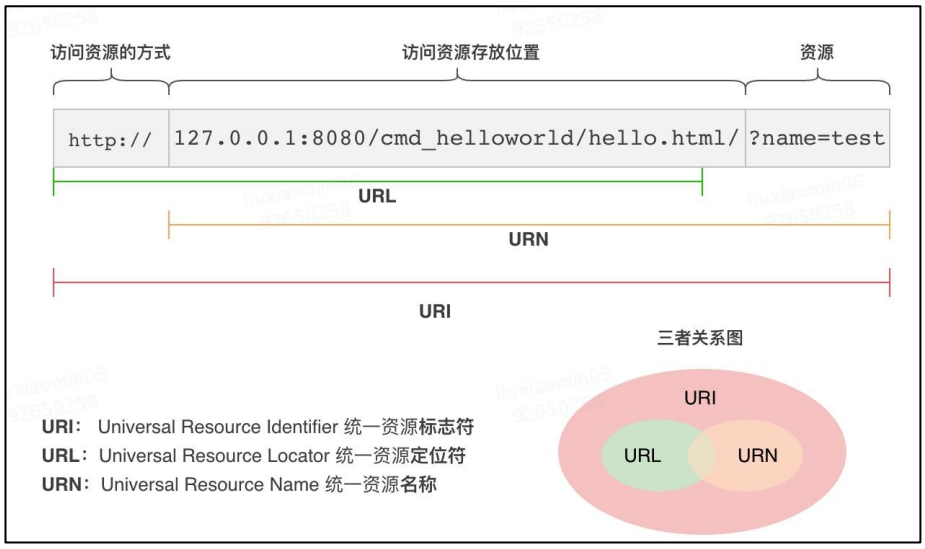


图 1：统一资源标识符架构

（3）身份验证

在现代社会体系下，身份的验证是信任建立的基础。当个体或组织之间发生交互关系时，均需要进行身份的验证，即证明某个体或组织拥有某项资源的所有权或享有某些权益，目的是通过身份验证系统维护系统运行的基本规则和安全。

① 物理世界身份验证

物理介质证明，如各种纸质文件或卡片证明，是人类发展史上依赖最长久的身份证明，包括身份证、护照、社会医疗保障卡、驾驶证等等。随着技术的发展，物理介质证明作假越来越容易，且在身份验证环节无法有效辨别，经常出现身份篡改、身份冒用等导致资产非法转移及社会权益盗用等问题。因此，通过物理介质实现身份证明来维护原有的社会规则和安全难以持续下去。为了防止身份作假，各政府及组织从两方面进行升级：第一方面是对身份证明的物理介质升级，增加了各种特征可供验证，如我国身份证上增加激光变色识别、增加微缩文字、视觉上呈现图层叠放等；这些升级只是增加了非法分子的作假成本，一旦他们掌握了这些技术，依然可以复制出各种身份证明，而无法从根本上杜绝作假问题；第二方面是提升验证手段，政府机构对接各类身份证明平台，能够在某主体享有权益或处置资产前，通过比对物理介质证明与系统信息进行身份真伪识别，这种模式下存在两方面问题，第一个问题是各类身份证明平台未全面联通，数据孤岛导致验证信息不完整；第二个问题是企业及其他个体用户无权对接身份识别平台，在日常交易合作中，无法通过该模式验证身份真伪。

② 互联网用户身份验证

在互联网世界中，身份验证主要依赖于用户名和密码。能够输入正确的信息，就意味着身份验证通过。这种验证体系存在两种问题，第一种是用户名和密码容易被网络攻击者盗用；第二是中心化平台对用户身份信息拥有绝对控制权，他们可以在未获得用户许可的情况下，根据自己的需要删除、增加、更改、甚至交易用户的身份信息。

2、身份信息安全问题

无论是物理世界还是互联网世界，都存在身份管理方面的问题，而且两个平行世界的身份证明逐渐融合。物理世界中的身份证明作假问题，借助互联网来加强身份验证能力；而互联网世界由于匿名信和不可追溯性导致的安全问题，通过与物理世界的身份映射方式来解决。我们解决了身份的真实性和可信性带来的困扰，但同时也给我们带来了新的麻烦，即身份的特征和行为暴露在网络中，被各个平台无视相关规定，肆意收集身份相关的行为信息并滥用这些信息。

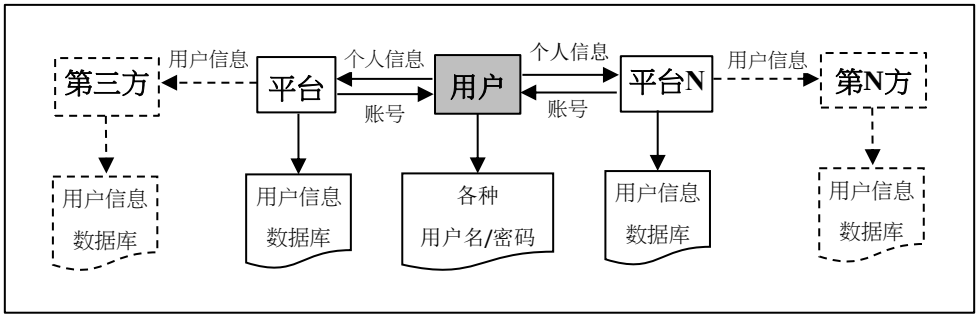


图 2：用户信息传统数据库管理模式

如图 2 所示，在中心化管理模式下，用户信息被不同平台重复收集并存储，在《从用户画像实现看数据隐私问题》（2021 年第 76 期）一文中，我们指出其中的问题，包括用户信息被过度采集、信息被不同平台交易、用户对个人行为数据没有控制权等问题。

3、其他

当前我们面对的不仅仅是上面所提到的关于人的身份管理问题，随着互联网技术及通信技术的发展，网络连接万事万物，构筑出一个与物理世界相平行的数字世界。数字世界里的参与者不仅仅是人，还有包括其他万事万物，如何定义数字世界里的这些万事万物所属权，以及怎么定义每一个数字对象的权益？这个问题关系着数字世界的正常秩序的维护以及信任的构建。前面提到的三要素“身份-身份证明-身份验证”仅围绕人来讨论，但物理世界中，除人的身份以外，我们还有其他各种国际统一标识，比如商品相关的统一编码（RFID，商品序列号，二维码）等。未来我们需要管理数字世界中每一种要素，前提是做好这些要素的身份管理。进一步说，我们需要一项能够统一维护不同身份标识方法的工具，能够做到不同事物的“身份标识-身份证明-身份验证”。

第二章 DID 技术详述

分布式数字身份（Decentralized Identifiers,简称 DIDs）,在 W3C 的《DID V1.0》中，将 DID 定义为一种新的全球唯一标识符。这种标识符不仅可以用于人，也可以用于万事万物，包括一辆车、一只动物，甚至是一台机器，本文主要以人为例来展开 DID 的讨论。

下面我们从技术实现和应用两个角度介绍 DID 技术，技术实现主要讲述 DID 技术的构

成要素；而应用主要围绕"身份-身份证明-身份验证"讨论 DID 的实现。

1、技术实现

DID 技术的核心构成要素包括三个：DID、DID Document 和 Verifiable Data Registry。

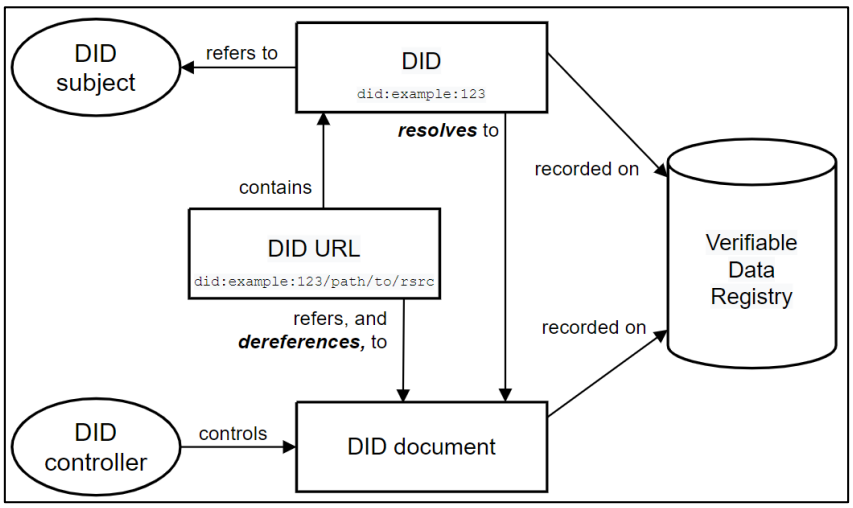


图 3：DID 架构及相关构成要素之间的关系

（资料来源：W3C DID core）

（1）DID

DID 属于统一资源标识符 URI 的一种，是一个永久不可变的字符串，它存在的意义有两点，第一，标记任何目标对象(DID Subject)，可以是一个人、一件商品、一台机器或者一只动物等等；第二，DID 是通过 DID URL 关联到描述目标对象的文件（DID Document,简称 DID Doc）唯一标识符，即通过 DID 能够在数据库中搜索到具体的 DID Doc。

① DID 标识方法

DID 分为三个部分，如图 4 所示，第一部分是 DID Scheme（类似 URL 中的 http,https,ftp 等协议）；第二部分是 DID 方法标识符（一般是 DID 方法的名称）；第三部分是 DID 方法中特定的标识符：在整个 DID 方法命名空间是唯一的。W3C 只规范了 DID 的表示结构，即 <did:+DID method:+DID Method-Specific Identifier>,但没有规范三部分内容的具体标准，具体内容与 DID Method 有关，将在下面第 2 部分介绍。

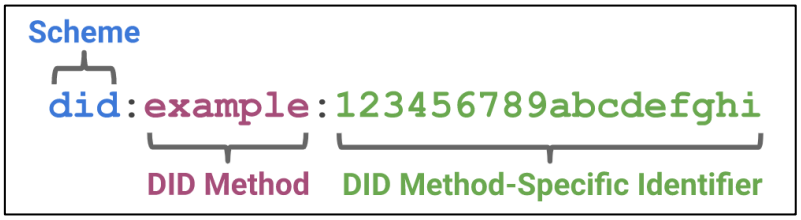


图 4：DID 简单示例

② DID Method

DID Method 是一组公开的操作标准，定义了 DID 的创建、解析、更新和删除，并涵盖了 DID 在身份系统中注册、替换、轮换、恢复和到期等。目前没有统一的操作标准，各个公司可以根据场景特征自行设计，由 W3C CCG 工作组统一维护。截至 2021 年 8 月 3 日发布《DID V1.0》,在 W3C 登记的 DID Method 高达 103 项，均有不同的名称和特定的标识符表示方法。

③ DID URL

为融合现有 URI 网络位置标识方法，DID 使用了 DID URL 表示资源的位置（如路径、查询和片段）。W3C 对 DID URL 的语法描述 ABNF 规定如下：<did-url = did path-abempty ["?"query]["#"fragment]>。

（2）DID Document

DID Document(DID Doc)包含着所有与 DID subject 有关的信息，在 Doc 中有身份信息验证方法（包括加密公钥，相关地址等）。DID Doc 是一个通用数据结构，通常是由 DID controller 负责数据写入和更改，文件里包含与 DID 验证相关的密钥信息和验证方法，提供了一组使 DID 控制者能够证明其对应 DID 控制的机制。需要说明的是，这里的管理 DID Doc 的 DID Controller 可能是 DID subject 本人，也有可能是第三方机构，不同 DID Method 对 DID Doc 的权限管理有所区别。

如图 5 所示，是一个与图 4 中的 DID 对应的 DID Doc（用 JSON-LD 编写的文件），存储在所有人能控制的位置（可以是中心化的，也可以去中心化的），以便轻松查找。文件中可能包含以下内容：

- 创建时间的时间戳记；
- DID Documents 有效的加密证明；
- 加密公钥列表；
- DID 可用于进行身份验证的方式列表

EXAMPLE 1: A simple DID document

```
{
  "@context": [
    "https://www.w3.org/ns/did/v1",
    "https://w3id.org/security/suites/ed25519-2020/v1"
  ]
  "id": "did:example:123456789abcdefghi",
  "authentication": [{
    // used to authenticate as did:...fghi
    "id": "did:example:123456789abcdefghi#keys-1",
    "type": "Ed25519VerificationKey2020",
    "controller": "did:example:123456789abcdefghi",
    "publicKeyMultibase": "zH3C2AVvLMv6gmMNam3uVAjZpfkcJCwDwnZn6z3wXmqPV"
  }]
}
```

图 5：DID Doc 示例

（3）Verifiable Data Registries(VDR)

DID 的初衷是将用户身份信息管理权从平台交回用户自己，这过程中用户必须解决的问

题是信息存储在哪里？以及需要验证的时候去哪里找到这些数据？怎么保证数据的真实性？VDR 讨论的就是怎么解决这些问题，我们将支持记录 DID 数据且能够在生成 DID Doc 时提供相关数据的系统称为 Verifiable Data Registry（VDR），这种系统包括分布式账本、分布式文件系统、P2P 网络或其他可被信任的渠道。目前市场主推 DID 储存媒介是钱包，分为托管钱包（如 Coinbase）、普通钱包（如 imtoken），以及智能钱包（Gnosis Safe, Dappe，Argent），具体哪种媒介能更有效的存储 DID 信息，暂不在本文详细讨论。

2、DID 的实现："身份-身份证明-身份验证"

我们基于第一部分的"身份-身份证明-身份验证"，简单讨论 DID 如何实现这些功能的。

（1）身份

在 DID 方案中，每个人可以在不同场景、不同时间，因为不同目的，在任意可信的第三平台登记不同的 DID，相关权益和资产所属权与不同的 DID 直接绑定，而身份主体通过持有 DID 来证明其对资产的所有权或具体权益。DID 没有直接与物理世界身份生成映射关系，且 DID 信息维护也是由身份主体或可信第三方来维护，保证了信息的安全性。对于身份主体而言，需要做好 DID 的安全持有工，同时维护好与 DID 对应的身份文件(DID Doc)。

（2）身份证明

DID 只是一串带有密钥的随机数值，在具体使用中第三方机构根据 DID 信息将身份证明写入 DID Doc，同时第三方机构会将自己的数字签名加入文件中，方便后期身份验证。例如，张三，需要证明自己具有驾驶能力，此时无需像传统的中心化方法，由权威机构签发一张驾驶证给张三，具体个人信息也无需存储在权威机构的数据库里； 通过 DID 技术给出的解决方案是[1]：张三向车管所提供自己准备的 DID 或使用车管所提供的 DID，车管所按照 DID Doc 的 JSON-LD 数据结构写入相关信息（包括但不限于 id，type,有效期，controller，验证方法等），同时加入车管所的数字签名。DID Doc 可以储存在车管所，可以储存在张三的智能钱包里，或者其他存储媒介。需要注意的是，此处 DID 并没有泄露张三的身份特征，没有映射物理世界的其他身份证明，这个 DID 只是张三持有的众多 DID 中的一个。因此，只要张三本人不出示 DID 证明，就没有人能知道这份 DID Doc 是张三的，从而保护了张三的个人隐私。

（3）身份验证

DID 验证方法有多种，具体方法在不同的 DID Method 方案中有所区别，较为常用的方法是零知识证明，例如根据国家最新青少年网络游戏规定中要求每天限时一小时，传统的方法则需要上传身份证信息，但分布式标识符解决方案中，只需要提供自己持有的 DID，通过零知识证明验证用户是否超过 18 岁即可，而无需告知平台方用户具体年龄。这只是众多验证方法中的一种。

[1]由于不同 DID Method 有不同操作建议，此处仅为举例说明 DID 身份证明签发的一种方法。

第三章 DID 的应用及发展

DID 从提出到现在已经有四年时间，各行业协会、互联网平台、基金会都在积极推动并完善 DID 技术。经过长时间探索，W3C 于 2021 年 8 月 3 日发布了 DID 1.0 版白皮书。相比初期 0.1 版搭建了一个全新的身份标识体系，到 1.0 版开始考虑如何融合市场上已有的身份标识方法。其他协会、组织及企业也基于 W3C 的 DID 规范提出了多种 DID Method,但距离 DID 技术落地应用，仍然有很多问题需要进一步去解决，主要包括：

1、如何满足合规性要求？

互联网最初只需要通过用户名/密码实现平台身份验证即可，但为了满足合规要求，增加了物理世界身份验证。这种方法初衷是为了让网络用户的行为可问责、可追溯，逐步建立网络信任体系，但负面影响是造成大量个人信息泄露。DID 有效解决了这些问题，但面临的仍然是合规性问题。虽然当前未出台相关规定，但不远将来肯定会面临如何将不同的 DID 映射到具体主体的问题，同时需要考虑这种映射关系，是否会造成新一轮信息泄露问题？该问题有待进一步探讨及观察。

2、如何验证 DID 与持有人之间的关系？

DID 具有匿名性，当前主流 DID 技术给出的解决方案是：谁持有 DID，谁就有权享受相关权益。这种方案无法验证 DID 提供者是不是本人，也无法避免 DID 被盗取并用于非法目的。虽然部分 DID Method 提出将 DID 映射到中心化数据库，通过中心化的一套方法验证 DID 提供者是不是本人的问题，但这仍将给个人信息保护留下漏洞，例如是否能够通过中心化数据库倒推出 DID 持有人？

3、DID 如何市场化推行？

DID 市场化过程中，当前有两方面瓶颈：第一方面，没有企业愿意主动放弃用户数据；用户数据如同平台护城河，产生了大量价值，如果同意 DID 的使用，就等于同意拆除护城河，对于互联网企业是致命性打击。第二方面，DID 技术推行谁来买单？第一，不同用户愿不愿意为自己的身份信息买单？换句话说，用户是否愿意向类似智能钱包这种供应方付费？虽然未来个人行为数据有机会变现，并足够支付这部分费用，但商业模式不清晰的情况下，有多少人感兴趣参与其中？第二，DID 技术将打破各平台方原有数据管理结构，必定需要新增相关验证平台，相关成本谁来承担？这些瓶颈将会极大阻碍 DID 技术的推行，如何平衡相关方利益关系，目前仍没有理想的方案。

4、密钥管理风险大

DID 的可信性主要依赖于密钥技术，如果第三方机构的私钥被窃取，会不会出现随意签发证书的行为？或者某个身份主体将私钥无意丢失，是否永远无法使用这些 DID 证书？这些问题目前没有非常理想的解决方案，对于现实使用也会提出较大挑战。

第四章 区块链与 DID 的结合

虽然区块链不是 DID 技术的必选项，但区块链技术能够助力与 DID 技术的实施，避免很多争端问题的发生，同时能够以更低成本维护数据的可信性，主要体现在以下几方面：

1、降低验证成本

DID 技术提倡的是将 DID 和 DID Doc 存储在用户端，但如何保证 DID Doc 不会在用户端被篡改？如果不使用区块链，则需要 DID 证书签发者同步维护这些证书，增加了维护成本。当使用 DID 时，合作方可以将 DID 持有人的 DID Doc 与签发方的数据库进行一致性验证，增加了验证成本。但使用区块链技术则能降低签发方这部分成本，只要写入持有人 DID Doc 的信息将被记录在链上，无法做出修改，保证了信息的真实性和安全性；签发人无需增加数据库存储及维护成本，而合作人也无需增加成本将持有人 DID Doc 与签发方数据库做一致性检验。

2、基于 DID 的信任体系的搭建

当前围绕区块链的方案大多没有实现生态闭环，假如有人在区块链生态中出现违约，仍需回归到中心化模式下寻找法律解决方法，并没有减轻政府治理压力。未来是否会通过构建基于 DID 的信用记录系统，来补齐这块生态闭环构建的短板？该问题值得观察。不同主体的行为信息会随着 DID 被记录在 DID Doc 中，也将成为不同相关方合作的重要参考，随着数据的增加违约的成本也将增加，例如无法从银行获取贷款、无法找到工作、无法找到合作伙伴，因此，这种信任体系将会对生态治理起到非常积极影响。而这一切的基础是可信数据，区块链不能缺席。

免责声明

本文件所载的信息由我们从被认为可靠的来源汇编而成，但我们、我们的附属公司或任何其他人士概不就该等信息的公平性、合理性、可靠性、准确性、完 整性或正确性作出任何明示或默示的声明或保证。本文件中包含的所有插图、示例或前瞻性信息（如有）仅在本文件发布之日出于说明目的而真诚提供，并 非旨在用作且不得被视为对事实或概率的担保、保证、预测或确定陈述而加以依赖。我们已努力确保在发布本文件时，文件所载信息的准确性和完整性，但 仍可能出现错误或遗漏。过往业绩无法说明未来表现，我们无法保证您未来能获得回报，而您可能会损失原始资本。我们保留更正任何错误或遗漏以及随时更改或更新信息的权利，恕不另行通知。

每个司法管辖区都有专门的法律，用以规管可能向其居民和 / 或在该管辖区范围内提供的投资和 / 或服务的类型，以及具体的规管流程。因此，本文件中讨 论的某些投资产品或服务可能无法在某些司法管辖区销售或供应。本文件并非出售要约或购买任何投资或服务的要约邀请。除非另有说明，否则我们概未表 示我们已获得在任何司法管辖区开展受规管活动的许可。此外，提供本材料并非且在任何情况下都不应被解释为，任何人士或公司在任何司法管辖区从事其 未获法律许可从事的受规管业务。

本文件中的任何内容都不构成法律、会计或税务建议，建议您在根据本文件中的任何内容采取行动之前，寻求独立的法律、税务和会计建议。本材料的内容 没有经过任何监管机构的审查；建议投资者对与本文件有关的任何投资或服务持谨慎态度。如果您对本材料的任何内容有疑问，您应该寻求得独立的专业意见。

在法律允许的最大范围内，HashKey Group 或任何其附属机构对因使用本文件或其中的信息而产生的任何直接或间接损失不会承担任何责任。未经 HashKey 的事先书面同意，本文件中的任何信息都不得以任何方式进行复制或抄袭。

"HashKey Group" 是一个品牌名称，用于描述由 HashKey 数字资产集团有限公司及其附属公司组成的集团公司中的任何一个或多个实体。

关于

HashKey Group 是亚洲领先的端到端数字资产管理和金融公司。HashKey 总部设于香港，也在新加坡和东京开展业务，并帮助机构投资者抓住数字资产和区块链技术的高潜力投资机会。该公司的核心业务包括一个计划中的数字资产交易所、全方位经纪平台、机构级托管服务和一个投向于全球区块链公司和数字资产项目的风险投资基金。

HashKey 提供区块链研究、开发人员和技术提升的机会。HashKey 还与领先的金融科技企业、学术机构和行业协会建立了广泛的合作关系。

HashKey 的高级领导团队拥有数十年的投资和交易经验、并具有强大的业务执行力和深刻的市场洞察力，致力为数字资产经济维持最高的合规和监管标准。

了解更多

www.hashkey.com
contact@hashkey.com