(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2019 年 6 月 6 日 (06.06.2019)

WIPO | PCT

(10) 国际公布号
WO 2019/105189 A1

(51) 国际专利分类号:
G06N 3/08 (2006.01)

(21) 国际申请号: PCT/CN2018/114082

(22) 国际申请日: 2018 年 11 月 6 日 (06.11.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201711227185.X 2017年11月29日 (29.11.2017) CN

(71) 申请人: 华 为 技 术 有 限 公 司 (**HUAWEI TECHNOLOGIES CO., LTD.**) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(72) 发明人: 陈 普 (**CHEN, Pu**); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 廖乔勃 (**LIAO, Qiaobo**); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

WO 2019/105189 A1

(54) **Title:** MODEL TRAINING SYSTEM AND METHOD, AND STORAGE MEDIUM

(54) 发明名称: 模型训练系统、方法和存储介质

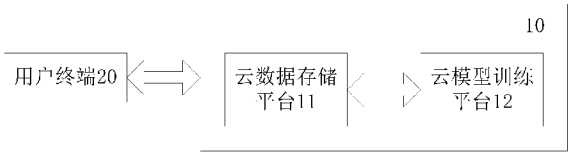

图 2

11 CLOUD DATA STORAGE PLATFORM
12 CLOUD MODEL TRAINING PLATFORM
20 USER TERMINAL

(57) **Abstract:** The present invention relates to the field of machine learning, and provides a model training system and method, and a storage medium. The model training system comprises a cloud data storage platform and a cloud model training platform. The cloud data storage platform is used for storing training data, receiving a training data call request, and exporting training data corresponding to a data call instruction to the cloud model training platform according to the training data call request. The cloud model training platform is used for receiving a model training creation instruction and obtaining a model to be trained, used for generating the training data call request and sending same to the cloud data storage platform, and used for training, by means of the training data exported from the cloud data storage platform, the model to be trained to obtain a training result model. The technical solution of the present invention can reduce the risk of training data leakage.

(57) 摘要: 本发明提供了一种模型训练系统、方法和存储介质，涉及机器学习领域。该模型训练系统，包括云数据存储平台和云模型训练平台；云数据存储平台用于存储训练数据，以及用于接收训练数据调用请求，根据训练数据调用请求，将与数据调用指令对应的训练数据导出至云模型训练平台；云模型训练平台用于接收模型训练创建指令，获取待训练模型，以及用于生成并向云数据存储平台发送训练数据调用请求，以及用于利用从云数据存储平台导出的训练数据，训练待训练模型，得到训练成果模型。利用本发明的技术方案能够降低训练数据发生泄露的风险。

[见续页]

(84) 指定国(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

模型训练系统、方法和存储介质

技术领域

本发明涉及机器学习领域，尤其涉及一种模型训练系统、方法和存储介质。

背景技术

深度学习广泛应用于人工智能和计算机视觉等领域。深度学习需要进行模型训练，在模型训练过程中模型开发者需要设计好特定模型，利用数据集进行多次迭代训练，从而得到符合期望要求的深度学习模型。其中，数据集是决定训练出的模型的稳定性和精确度是否符合期望要求的关键。数据集可由数据提供者提供。

现阶段，用户可在数据提供商处购买下载数据权限。下载数据权限通过后，用户可将数据下载至本地保存。当需要进行模型训练时，将下载至本地保存的数据拷贝到模型训练系统中，实现模型训练。但是，下载至本地保存的数据发生泄漏的风险较大。

发明内容

本申请提供了一种模型训练系统、方法和存储介质，能够降低训练数据发生泄露的风险。

第一方面，本申请提供了一种模型训练系统，包括云数据存储平台和云模型训练平台；云数据存储平台用于存储训练数据，以及用于接收训练数据调用请求，根据训练数据调用请求，将与数据调用指令对应的训练数据导出至云模型训练平台；云模型训练平台用于接收模型训练创建指令，获取待训练模型，以及用于生成并向云数据存储平台发送训练数据调用请求，以及用于利用从云数据存储平台导出的训练数据，训练待训练模型，得到训练成果模型。

根据第一方面，在第一方面的第一种可能中，模型训练系统还包括检索数据平台和鉴权中心；云数据存储平台包括权限网关；检索数据平台用于根据数据提供者提供的训练数据，建立数据索引表，以及用于接收检索指令，根据检索指令在数据索引表中进行数据检索，并生成检索结果，以及用于接收用户终端针对检索结果的数据选取指令，根据数据选取指令向鉴权中心发起鉴权许可请求，鉴权许可请求包括训练数据的数据标识；鉴权中心用于接收鉴权许可请求，根据鉴权许可请求创建数据标识的数据令牌，并将数据令牌下发给权限网关和用户终端；云模型训练平台还用于向权限网关发送训练数据调用请求，训练数据调用请求包括鉴权中心下发至用户终端的数据令牌；权限网关用于建立第一对应关系，第一对应关系为数据标识与数据令牌一一对应的关系，以及用于接收训练数据调用请求，根据训练数据调用请求中的数据令牌，在第一对应关系中查找目标数据标识，目标数据标识为与训练数据调用请求中的数据令牌对应的数据标识，以及用于将目标数据标识对应的训练数据导出至云模型训练平台。

根据第一方面，在第一方面的第二种可能中，模型训练系统还包括检索数据平台和鉴权中心；云数据存储平台包括权限网关和至少一个数据存储服务器；检索数据平台用于根据数据提供者提供的训练数据，建立数据索引表，以及接收检索指令，根据检索指令在数据索引表中进行数据检索，并生成检索结果，以及用于接收用户终端针对检索结果的数据选取指令，根据数据选取指令向鉴权中心发起鉴权许可请求，鉴权许可请求包括训练数据的数据标识；鉴权中心用于接收鉴权许可请求，根据鉴权许可请求创建数据标识的数据令牌，并将数据令牌下发给权限网关和用户终端；云模型训练平台还用于向权限网关发送训练数据调用请求，训练数据调用请求包括鉴权中心下发至用户终端的数据令牌；权限网关用于建立第二对应关系，第二对应关系为数据令牌与数据路由的对应关系，数据路由包括训练数据的统一资源定位符路径，以及用于接收训练数据调用请求，根据训练数据调用请求中的数据令牌，在第二对应关系中查找目标数据路由，目标数据路由为与训练数据调用请求中的数据令牌对应的数据路由，以及用于访问目标数据存储服务器，以将目标数据存储服务器中目标数据路由指示的训练数据导出至云模型训练平台，目标数据存储服务器为与目标数据路由对应的数据存储服务器。

根据第一方面的第二种可能，在第一方面的第三种可能中，模型训练系统还包括访问路由器，权限网关通过访问路由器中预定的标准访问接口从目标数据存储服务器中导出目标数据路由指示的训练数据。

根据第一方面的第一种可能或第二种可能，在第一方面的第四种可能中，权限网关还用于获取更新判断参数，判断更新判断参数是否满足更新条件，以及用于若判定更新判断参数满足更新条件，向鉴权中心发送更新请求，以及用于与鉴权中心同步更新数据令牌；鉴权中心还用于接收更新请求，根据更新请求更新数据令牌。

根据第一方面的第四种可能，在第一方面的第五种可能中，更新判断参数包括对鉴权许可请求的拒绝次数；权限网关还用于监测鉴权中心对鉴权许可请求的处理过程，以及用于若监测到鉴权中心对鉴权许可请求的拒绝次数超出更新条件中的拒绝次数更新阈值，则向鉴权中心发送更新请求。

根据第一方面的第五种可能，在第一方面的第六种可能中，更新判断参数包括训练数据的调用次数；权限网关还用于获取一段时长内的训练数据的调用次数，以及用于若在一段时长内，同一训练数据的调用次数超出更新条件中的调用次数更新阈值，则向鉴权中心发送更新请求。

根据第一方面，在第一方面的第七种可能中，云模型训练平台还用于训练得到训练成果模型后，销毁云模型训练平台内训练训练成果模型所利用的训练数据和待训练模型。

根据第一方面，在第一方面的第八种可能中，模型训练系统还包括数据稽查系统；数据稽查系统用于对数据提供者上传的训练数据进行有效性认证，拒绝将有效性认证失败的训练数据存入云数据存储平台。

根据第一方面，在第一方面的第九种可能中，模型训练系统还包括云模型存储平台；云模型存储平台用于提供待训练模型，以及保存训练成果模型。

根据第一方面的第九种可能，在第一方面的第十种可能中，模型训练系统还包括镜像平台和模型推理平台；镜像平台用于存储模型推理运行环境；模型推理平台用于接收推理请求，推理请求包括待处理数据，以及从镜像平台加载模型推理运行环境，以及从云模型

存储平台调用训练成果模型，将待处理数据导入训练成果模型进行模型推理。

第二方面，本申请提供了一种模型训练方法，包括：云模型训练平台接收模型训练创建指令，获取待训练模型；云模型训练平台生成并向云数据存储平台发出训练数据调用请求，以调用云数据存储平台中存储的训练数据；云数据存储平台接收训练数据调用请求，将与训练数据调用请求对应的训练数据导出至云模型训练平台；云模型训练平台利用从云数据存储平台导出的训练数据，训练待训练模型，得到训练成果模型。

根据第二方面，在第二方面的第一种可能中，上述模型训练方法还包括：检索数据平台根据数据提供者提供的训练数据，建立数据索引表；检索数据平台接收检索指令，根据检索指令在数据索引表中进行数据检索，并生成检索结果；检索数据平台接收用户终端的数据选取指令，根据数据选取指令向鉴权中心发起鉴权许可请求，鉴权许可请求包括训练数据的数据标识；鉴权中心接收鉴权许可请求，根据鉴权许可请求创建数据标识的数据令牌，并将数据令牌下发给权限网关和用户终端；权限网关根据下发得到的数据令牌，建立第一对应关系，第一对应关系为数据标识与数据令牌一一对应的关系。

根据第二方面的第一种可能，在第二方面的第二种可能中，云模型训练平台生成并向云数据存储平台发送训练数据调用请求，包括：云模型训练平台生成并向权限网关发送训练数据调用请求，训练数据调用请求包括鉴权中心下发至用户终端的数据令牌；云数据存储平台接收训练数据调用请求，将与训练数据调用请求对应的训练数据导出至云模型训练平台，包括：云数据存储平台中的权限网关接收训练数据调用请求，根据训练数据调用请求中的数据令牌，在第一对应关系中查找目标数据标识，并将目标数据标识对应的训练数据导出至云模型训练平台，目标数据标识为与训练数据调用请求中的数据令牌对应的数据标识。

根据第二方面，在第二方面的第三种可能中，上述模型训练方法还包括：检索数据平台根据数据提供者提供的训练数据，建立数据索引表；检索数据平台接收检索指令，根据检索指令在数据索引表中进行数据检索，并生成并发送检索结果；检索数据平台接收用户终端针对检索结果的数据选取指令，根据数据选取指令向鉴权中心发起鉴权许可请求，鉴权许可请求包括训练数据的数据标识；鉴权中心接收鉴权许可请求，根据鉴权许可请求创建数据标识的数据令牌，并将数据令牌下发给权限网关和用户终端；权限网关根据下发得到的数据令牌，建立第二对应关系，第二对应关系为数据令牌与数据路由的对应关系，数据路由包括训练数据的统一资源定位符路径。

根据第二方面的第三种可能，在第二方面的第四种可能中，云模型训练平台生成并向云数据存储平台发送训练数据调用请求，包括：云模型训练平台生成并向权限网关发送训练数据调用请求，训练数据调用请求包括鉴权中心下发至用户终端的数据令牌；云数据存储平台接收训练数据调用请求，将与训练数据调用请求对应的训练数据导出至云模型训练平台，包括：云数据存储平台中的权限网关接收训练数据调用请求，根据训练数据调用请求中的数据令牌，在第二对应关系中查找目标数据路由，目标数据路由为与训练数据调用请求中的数据令牌对应的数据路由；权限网关访问目标数据存储服务器，以将目标数据存储服务器中目标数据路由指示的训练数据导出至云模型训练平台，目标数据存储服务器为与目标数据路由对应的数据存储服务器。

根据第二方面或第二方面的第一种可能至第四种可能中的任意一种可能，在第二方面的第五种可能中，上述模型训练方法还包括：权限网关获取更新判断参数，判断更新判断参数是否满足更新条件；若判定更新判断参数满足更新条件，权限网关向鉴权中心发送更新请求；鉴权中心接收更新请求，根据更新请求更新数据令牌；权限网关与鉴权中心同步更新数据令牌。

根据第二方面的第五种可能，在第二方面的第六种可能中，更新判断参数包括对鉴权许可请求的拒绝次数；权限网关获取更新判断参数，判断更新判断参数是否满足更新条件，包括：权限网关监测鉴权中心对鉴权许可请求的处理过程，并获取鉴权中心对鉴权许可请求的拒绝次数，并判断鉴权中心对鉴权许可请求的拒绝次数是否超出更新条件中的拒绝次数更新阈值；若判定更新判断参数满足更新条件，权限网关向鉴权中心发送更新请求，包括：若监测到鉴权中心对鉴权许可请求的拒绝次数超出更新条件中的拒绝次数更新阈值，则向鉴权中心发送更新请求。

根据第二方面的第五种可能，在第二方面的第七种可能中，更新判断参数包括训练数据的调用次数；权限网关获取更新判断参数，判断更新判断参数是否满足更新条件，包括：权限网关获取一段时长内的训练数据的调用次数，判断在一段时长内，同一训练数据的调用次数是否超出更新条件中的调用次数更新阈值；若判定更新判断参数满足更新条件，权限网关向鉴权中心发送更新请求，包括：若在一段时长内，同一训练数据的调用次数超出更新条件中的调用次数更新阈值，则向鉴权中心发送更新请求。

根据第二方面，在第二方面的第八种可能中，在云模型训练平台利用从云数据存储平台导出的训练数据，训练待训练模型，得到训练成果模型之后，还包括：云模型训练平台销毁云模型训练平台内训练训练成果模型所利用的训练数据和待训练模型。

根据第二方面，在第二方面的第九种可能中，上述模型训练方法还包括：数据稽查系统对数据提供者上传的训练数据进行有效性认证；数据稽查系统拒绝将有效性认证失败的训练数据存入云数据存储平台。

根据第二方面，在第二方面的第十种可能中，在云模型训练平台利用从云数据存储平台导出的训练数据，训练待训练模型，得到训练成果模型之后，还包括：云模型存储平台保存训练成果模型。

根据第二方面的第十种可能，在第二方面的第十一种可能中，上述模型训练方法还包括：模型推理平台接收推理请求，推理请求包括待处理数据；模型推理平台从镜像平台加载模型推理运行环境，并从云模型存储平台调用训练成果模型，将待处理数据导入训练成果模型进行模型推理。

第三方面，本申请提供了一种存储介质，存储介质上存储有程序，程序被处理器执行时实现上述技术方案中的模型训练方法。

本申请提供了一种模型训练系统、方法和存储介质，可应用于深度学习场景中。模型训练系统可包括云数据存储平台和云模型训练平台。云数据存储平台存储训练数据。云模型训练平台接收用户的模型训练创建指令，触发执行模型训练。云模型训练平台通过向云数据存储平台发送训练数据调用请求，调用云数据存储平台存储的训练数据。云模型训练平台利用获取的待训练模型和从云数据存储平台导出的训练数据进行模型训练。在本申请中，云数据存储平台和云模型训练平台相互独立，将训练数据的存储与模型训练两种功能

分离。云数据存储平台和云模型训练平台均以云系统为基础实现，模型训练过程在云系统中进行，进行模型训练的用户无法将训练数据下载至本地，训练数据存在于云数据存储平台和正在进行模型训练的云模型训练平台。也就是说，训练数据不会从本地的用户侧泄露，从而降低了训练数据发生泄露的风险。

附图说明

图 1 为本发明实施例的模型训练系统的应用场景示意图；

图 2 为本发明一实施例中一种模型训练系统的结构示意图；

图 3 为本发明另一实施例中一种模型训练系统的结构示意图；

图 4 为本发明又一实施例中一种模型训练系统的结构示意图；

图 5 为本发明一实施例中一种模型训练方法的流程图；

图 6 为本发明一实施例中一种模型训练方法的一种具体实现方式的流程图；

图 7 为本发明一实施例中一种模型训练方法的另一种具体实现方式的流程图。

具体实施方式

本发明实施例提供一种模型训练系统、方法和存储介质，可应用于深度学习（Deep Learning）的场景中，可实现对深度学习模型的训练，也可实现对深度学习模型的应用，比如，利用训练处的深度学习模型进行推理。本发明实施例的模型训练系统可在云端完成模型训练、模型推理等功能。图 1 为本发明实施例的模型训练系统的应用场景示意图。如图 1 所示，模型训练系统可在云服务系统上运行，云服务系统可由云系统以及向外提供访问接口的系统集群网关构成。用户可通过用户终端使用账号及密码通过网络连接到云系统。云系统包括多个内部网络互通的服务器。模型训练系统可通过数据模型仓库实现训练数据和训练模型的存储和提供。模型训练系统可通过深度学习数据库实现模型训练系统与用户的人机交互，可通过鉴权服务系统完成用户与模型训练系统的各项权利的鉴权，可通过训练推理系统完成模型的训练和推理。

图 2 为本发明一实施例中一种模型训练系统的结构示意图。如图 2 所示，模型训练系统包括云数据存储平台 11 和云模型训练平台 12。

云数据存储平台 11 用于存储训练数据，以及用于接收训练数据调用请求，根据训练数据调用请求，将与数据调用指令对应的训练数据导出至云模型训练平台 12。

训练数据为用于对训练模型所需的数据，云数据存储平台 11 可存储多个训练数据，训练数据可视为由多条数据形成的数据集。训练数据可包括图像、视频、音频等，在此并不限定。云数据存储平台 11 在存储训练数据时，可为训练数据分配数据标识，数据标识用于标识训练数据，可作为查找数据存储位置的标识符。在一个示例中，为了区分不同的训练数据，训练数据的数据标识具有唯一性，也就是说，不同的训练数据的数据标识不同。

云数据存储平台 11 可接收数据提供者上传的训练数据。示例性地，数据提供者可利用客户端通过超文本传输协议（HyperText Transfer Protocol，HTTP）连接到云系统的后端，从而与云数据存储平台 11 进行信息交互。在一个示例中，云数据存储平台 11 可向数据提供者提供上传训练数据的标准协议，标准协议中可包括数据格式、压缩格式以及数据类型等。云数据存储平台 11 可对数据提供者上传的训练数据进行检测，若确定数据提供者上

传的训练数据不符合标准协议，则云数据存储平台 11 可拒绝存储不符合标准协议的训练数据。

云数据存储平台 11 中可设置一备份区域，该备份区域可用于对训练数据进行备份，避免数据出现意外，如数据误操作等导致无法恢复的情况。

训练数据调用请求是云模型训练平台 12 生成并发送的，根据训练数据调用请求可得知云模型训练平台 12 请求调用的训练数据。在一个示例中，训练数据调用请求可包括数据标识。云数据存储平台 11 接收训练数据调用请求，可查找训练数据调用请求需要调用的训练数据，并将请求调用的训练数据导出至云模型训练平台 12，以供云模型训练平台 12 利用导出的训练数据进行模型训练。

云模型训练平台 12 用于接收模型训练创建指令，获取待训练模型，以及用于生成并向云数据存储平台 11 发送训练数据调用请求，以及用于利用从云数据存储平台 11 导出的训练数据，训练待训练模型，得到训练成果模型。

其中，云模型训练平台 12 可获取用户或模型提供者上传的待训练模型，也可从云系统中的模型数据库中获取待训练模型。

在一个示例中，示例性地，用户可利用用户终端 20 通过超文本传输协议连接到云系统的后端，从而与云模型训练平台 12 进行信息交互。用户可通过用户终端 20 向云模型训练平台 12 发送模型训练创建请求，以触发云模型训练平台 12 创建模型训练任务。云模型训练平台 12 可利用待训练模型和训练数据进行模型训练。示例性的，模型训练可指将训练数据导入待训练模型进行多次迭代训练，从而得到经训练后的模型即训练成果模型。

需要说明的是用户终端 20 的使用者可包括用户、数据提供者或模型提供者。

本发明实施例中的云数据存储平台 11 可视为图 1 中数据模型仓库的一部分。本发明实施例中的云模型训练平台 12 可视为图 1 中训练推理系统的一部分。

在本发明实施例中，云数据存储平台 11 和云模型训练平台 12 相互独立，将训练数据的存储与模型训练两种功能分离。云数据存储平台 11 和云模型训练平台 12 均以云系统为基础实现，模型训练过程在云系统中进行，进行模型训练的用户无法将训练数据下载至本地，训练数据存在于云数据存储平台 11 和正在进行模型训练的云模型训练平台 12。也就是说，训练数据不会从本地的用户侧泄露，从而降低了训练数据发生泄露的风险。

图 3 为本发明另一实施例中一种模型训练系统的结构示意图。图 3 与图 2 的不同之处在于，图 2 中的云数据存储平台 11 还包括图 3 中的权限网关 111；图 3 所示的模型训练系统还可包括检索数据平台 13、鉴权中心 14、数据稽查系统 15、云模型存储平台 16、镜像平台 17 和模型推理平台 18。

检索数据平台 13 用于根据数据提供者提供的训练数据，建立数据索引表。用户可通过检索数据平台 13 对云数据存储平台 11 中存储的训练数据进行搜索查询。

在一个示例中，在数据提供者上传训练数据后，检索数据平台 13 可对训练数据进行分析处理，得到训练数据的数据集大小、数据集规模、数据所有者信息、数据上传日期等数据基本信息，便于用户了解训练数据的基本信息。

在一个示例中，云数据存储平台 11 还可要求数据提供者在上传训练数据时，提供训练数据的标签，训练数据的标签可表征训练数据的特征。具体的，训练数据的标签可以为训练数据表征的内容的关键词。比如，数据提供者在上传训练数据时，为训练数据标记的

标签为“车牌”和“小型车”。检索数据平台 13 在建立数据索引表的过程中，也可将训练数据的标签添加入数据检索表，以便于用户在检索训练数据时，利用训练数据的特征进行检索。

检索数据平台 13 用于接收检索指令，根据检索指令在数据索引表中进行数据检索，并生成检索结果。具体的，检索指令中可包括一个或多个检索关键词，可根据检索关键词在数据索引表中的训练数据的标签中进行查找。检索结果可包括与检索指令中的检索关键词相关的训练数据的信息，比如训练数据的名称、编号、关键词以及训练数据中的部分数据示例等。在一个示例中，检索结果可包括按照与检索关键词的相关程度的大小依次排列的训练数据的信息，使用户能够更直观地得到与检索关键字最相关的训练数据。在另一个示例中，也可在根据检索关键词检索到的训练数据的信息中随机筛选固定数据的训练数据的信息提供给用户。比如，每次检索生成的检索结果包括十条训练数据的信息。检索数据平台 13 可将检索结果发送给用户终端 20，用户终端 20 可显示检索结果。

用户接收到检索结果后，还可通过用户终端 20 针对检索结果发出数据选取指令。数据检索平台接收用户终端 20 针对检索结果的数据选取指令，根据数据选取指令向鉴权中心 14 发起鉴权许可请求。数据选取指令可用于指示选取检索结果中的一项或多项训练数据的信息，从而确定模型训练需要的训练数据。

确定模型训练需要的训练数据后，向鉴权中心 14 发起鉴权许可请求，鉴权许可请求可包括训练数据的数据标识，向鉴权中心 14 请求训练数据的调用权限。

本发明实施例中的检索数据平台 13 可视为图 1 中的深度学习数据库的至少一部分。

鉴权中心 14 用于接收鉴权许可请求，根据鉴权许可请求创建数据标识的数据令牌，并将数据令牌下发给权限网关 111 和用户终端 20。

鉴权许可请求用于请求训练数据的调用权限。鉴权中心 14 可决定是否同意检索数据平台 13 发送来的鉴权许可请求。示例性的，鉴权许可请求可包括针对训练数据的付费信息，若付费信息表明用户对针对训练数据付费成功，鉴权中心 14 可同意鉴权许可请求，并创建数据标识的数据令牌。鉴权中心 14 同意鉴权许可请求后，还可生成并保存数据鉴权信息，数据鉴权信息可包括用户标识和数据标识。示例性的，数据鉴权信息可具有有效时长，即在有效时长内，若用户再次请求同样的训练数据时，鉴权许可请求可直接被鉴权中心 14 同意通过，不需要进行审核。有效时长可根据工作场景和工作需求设定，在此并不限定。比如，有效时长可为一年或永久。

数据令牌（即数据 Token）可标识某个操作中的训练数据，作为数据调用的一种安全凭证使用。比如，数据令牌标识后续过程中数据调用操作中的训练数据。在一个示例中，数据令牌可实现为安全插件。鉴权中心 14 将创建的数据令牌下发给用户终端 20，以使得用户终端 20 可利用数据令牌通过权限网关 111 从云数据存储平台 11 导出与数据令牌对应的训练数据。同时，鉴权中心 14 也将创建的数据令牌保存在鉴权中心 14。

云模型训练平台 12 还用于向权限网关 111 发送训练数据调用请求，训练数据调用请求包括鉴权中心 14 下发至用户终端 20 的数据令牌。

比如，用户终端 20 在请求训练数据时，可将数据令牌添加入模型训练创建指令，云模型训练平台 12 可解析模型训练创建指令，得到下发至用户终端 20 的数据令牌，并将下发至用户终端 20 的数据令牌添加入训练数据调用请求中。云模型训练平台 12 通过训练数

据调用请求中的数据令牌从云数据存储平台 11 调用与数据令牌对应的训练数据。

在一种实现方式中，云数据存储平台 11 具体可实现为第三方公用服务器。第三方公用服务器不属于数据提供者、模型提供者和用户，是一个公用的用于存储训练数据且能够导出训练数据的服务器。调用训练数据可利用数据令牌与数据标识的对应关系进行授权调用。

权限网关 111 用于建立第一对应关系，第一对应关系为数据标识与数据令牌的对应关系。数据标识与数据令牌一一对应，数据令牌也具有唯一性，也就是说，不同的数据标识对应不同的数据令牌。权限网关 111 在接收到训练数据调用请求时，根据训练数据调用请求中的数据令牌，在第一对应关系中查找目标数据标识，目标数据标识为与训练数据调用请求中的数据令牌对应的数据标识，并将目标数据标识对应的训练数据导出至云模型训练平台 12。

当云数据存储平台 11 接收到训练数据调用请求后，权限网关 111 会对比训练数据调用请求中的数据令牌是否与权限网关 111 中存储的数据令牌；若训练数据调用请求中的数据令牌能够与权限网关 111 中存储的数据令牌匹配，则允许调用训练数据，并将与训练数据调用请求中的数据令牌对应的训练数据导出。

为了保障模型训练过程中的数据安全，避免训练数据被越权使用，可根据实际情况对数据令牌进行更新。权限网关 111 可用于获取更新判断参数，判断更新判断参数是否满足更新条件。若判定更新判断参数满足更新条件，权限网关 111 向鉴权中心 14 发送更新请求，以及用于与鉴权中心 14 同步更新数据令牌。鉴权中心 14 接收更新请求，根据更新请求更新数据令牌。

更新判断参数可包括对鉴权许可请求的拒绝次数、训练数据的调用次数、数据令牌的存在时长等参数中的一项或多项。

比如，更新判断参数包括对鉴权许可请求的拒绝次数。权限网关 111 可监测鉴权中心 14 对鉴权许可请求的处理过程，从而得到鉴权中心 14 对鉴权许可请求的拒绝次数。若权限网关 111 监测到鉴权中心 14 对鉴权许可请求的拒绝次数超出更新条件中的拒绝次数更新阈值，则向鉴权中心 14 发送更新请求。

拒绝次数更新阈值可根据工作场景和工作需求设定，在此并不限定。鉴权中心 14 删除原数据令牌，并生成新的数据令牌，并将新的数据令牌下发给客户终端和权限网关 111，以使得权限网关 111 可以与鉴权中心 14 同步更新数据令牌。数据令牌在鉴权中心 14 和权限网关 111 中更新时，需要停止训练数据调用请求的执行，待鉴权中心 14 和权限网关 111 中的数据令牌更新完毕后，再执行训练数据调用请求。在数据令牌更新完毕后，若训练数据调用请求中包含的仍然是原数据令牌，训练数据调用请求中的原数据令牌失效，则无法调用训练数据。

又比如，更新判断参数包括训练数据的调用次数。权限网关 111 可获取一段时长内的训练数据的调用次数。若在一段时长内，权限网关 111 确定同一训练数据的调用次数超出更新条件中的调用次数更新阈值，则向鉴权中心 14 发送更新请求。统计训练数据的一段时长和调用次数更新阈值可根据工作场景和工作需求设定，在此并不限定。

还比如，更新判断参数包括数据令牌的存在时长。权限网关 111 可设置数据令牌的更新周期时长，并记录数据令牌的存在时长。若权限网关 111 确定数据令牌的存在时长达到

更新周期时长，则向鉴权中心 14 发送更新请求。数据令牌的更新周期时长可根据工作场景和工作需求设定，在此并不限定。

需要说明的是，更新判断参数和更新条件并不限于上述举例。权限网关 111 也可接收用户的更新策略配置指令，根据更新策略配置指令设置更新判断参数和更新条件。

云模型存储平台 16 用于提供待训练模型，以及保存训练成果模型。云模型存储平台 16 中存储的模型可以是模型提供者上传的模型，也可以是云模型训练平台 12 训练得到的训练成果模型。

在一个示例中，上述云模型训练平台 12 在训练得到训练成果模型后，可将训练成果模型发送至云模型存储平台 16 保存，并销毁云模型训练平台 12 内训练训练成果模型所利用的训练数据和待训练模型，还可将云模型训练平台 12 内的训练成果模型销毁，以防止遗留在云模型训练平台 12 的训练数据和模型即待训练模型和训练成果模型泄露。

在一个示例中，数据稽查系统 15 先于云数据存储平台 11 接收到数据提供者上传的训练数据。数据稽查系统 15 用于对数据提供者上传的训练数据进行有效性认证，拒绝将有效性认证失败的训练数据存入云数据存储平台 11。比如，若数据提供者上传的训练数据与云数据存储平台 11 存储的训练数据重复，或者数据提供者上传的数据的数据格式不符合云数据存储平台 11 的标准协议，则数据稽查系统 15 判定数据提供者上传的训练数据无效，即上传的训练数据有效性认证失败。若数据稽查系统 15 判定数据提供者上传的训练数据有效，则可通过检索数据平台 13 向云数据存储平台 11 发送存储指令，以使得云数据存储平台 11 将数据提供者上传的训练数据持久存储。

需要说明的是，对数据提供者上传的训练数据进行有效性认证的方式并不限于上述方式。数据稽查系统 15 可保证模型训练系统中所使用的训练数据的真实有效性。

镜像平台 17 用于存储模型推理运行环境。具体的，模型推理运行环境可包括系统环境和训练成果模型对应的运行框架环境。

模型推理平台 18 可接收推理请求，推理请求包括待处理数据。推理请求可由用户终端 20 发送。示例性的，用户终端 20 可通过应用程序编程接口（Application Programming Interface，API）向模型推理平台 18 发送推理请求。模型推理平台 18 接收推理请求后，从镜像平台 17 加载模型推理运行环境，并从云模型存储平台 16 调用训练成果模型，将待处理数据导入训练成果模型进行模型推理。

本发明实施例中的数据检索平台可视为图 1 中深度学习数据库中的至少一部分。本发明实施例中的鉴权中心 14 可视为图 1 中鉴权服务系统中的至少一部分。本发明实施例中的模型推理平台 18 可视为图 1 中训练推理系统中的一部分。

图 4 为本发明又一实施例中一种模型训练系统的结构示意图。图 4 所示的模型训练系统与图 3 所示的模型训练系统的不同之处在于，云数据存储平台 11 可实现为数据提供者的至少一个私有服务器。

在云数据存储平台 11 包括权限网关 111 和至少一个数据存储服务器 112 即私有服务器的条件下，调用训练数据可利用数据令牌与数据路由的对应关系进行授权调用。

数据路由可包括训练数据的统一资源定位符（Uniform Resource Locator，URL）路径，还可包括数据访问方法和从云数据存储平台 11 导出训练数据的标准。数据提供者在上传训练数据的同时也可上传训练数据对应的数据路由至检索数据平台 13。

检索数据平台 13 也可对数据路由进行合法性检测，若确定数据路由不合法，则拒绝存储数据路由。比如，检索数据平台 13 确定数据路由无法访问或数据路由的格式不符合模型训练系统中预设的标准，则拒绝存储数据路由。示例性的，检索数据平台 13 可向权限网关 111 和鉴权中心 14 发送拒绝指令，以使得权限网关 111 和鉴权中心 14 均拒绝存储路由数据。

权限网关 111 可建立第二对应关系，第二对应关系为数据令牌与数据路由的对应关系。示例性的，第二对应关系可实现为数据路由表。训练数据具有对应的数据路由，训练数据与数据令牌一一对应，数据令牌与数据路由也一一对应。在检索数据平台 13 建立数据索引表时，可将对应的数据路由保存在权限网关 111 中。

权限网关 111 接收训练数据调用请求后，根据训练数据调用请求中的数据令牌，在第二对应关系中查找目标数据路由。目标数据路由为与训练数据调用请求中的数据令牌对应的数据路由。权限网关 111 可根据与数据令牌对应的数据路由，访问目标数据存储服务器 112，以将目标数据存储服务器 112 中目标数据路由指示的训练数据导出至云模型训练平台 12。目标数据存储服务器 112 为与目标数据路由对应的数据存储服务器 112。

为了保证数据存储服务器 112 即私有服务器中的训练数据的安全性，可建立安全加密远程访问。在一个实例中，模型训练系统还可包括访问路由器。权限网关 111 通过访问路由器中预定的标准访问接口从目标数据存储服务器 112 中导出目标数据路由指示的训练数据。比如，标准访问接口为 restful 访问接口，并可将 restful 访问接口的路径作为数据路由。

在一个示例中，为了进一步保证数据存储服务器 112 中的训练数据的安全性。权限网关 111 可随机选取数据令牌，并验证数据令牌的合法性。若权限网关 111 确定数据令牌非法，则可更新数据路由表，即更新第二对应关系，具体可实现为更新第二对应关系中的数据令牌。

图 5 为本发明一实施例中一种模型训练方法的流程图。该模型训练方法可适用于上述实施例中的模型训练系统。如图 5 所示，模型训练方法可包括步骤 S201 和步骤 S204。

在步骤 S201 中，云模型训练平台接收模型训练创建指令，获取待训练模型；

在步骤 S202 中，云模型训练平台生成并向云数据存储平台发出训练数据调用请求，以调用云数据存储平台中存储的训练数据；

在步骤 S203 中，云数据存储平台接收训练数据调用请求，将与训练数据调用请求对应的训练数据导出至云模型训练平台；

在步骤 S204 中，云模型训练平台利用从云数据存储平台导出的训练数据，训练待训练模型，得到训练成果模型。

上述步骤 S201 至步骤 S204 的说明可参见上述实施例中的云模型训练平台和云数据存储平台的相关说明。

在本发明实施例中，云数据存储平台和云模型训练平台相互独立，将训练数据的存储与模型训练两种功能分离。云数据存储平台和云模型训练平台均以云系统为基础实现，模型训练过程在云系统中进行，进行模型训练的用户无法将训练数据下载至本地，训练数据存在于云数据存储平台和正在进行模型训练的云模型训练平台。也就是说，训练数据不会从本地的用户侧泄露，从而降低了训练数据发生泄露的风险。

图 6 为本发明一实施例中一种模型训练方法的一种具体实现方式的流程图。如图 6 所示，模型训练方法可包括步骤 S301 至步骤 S315。

在步骤 301 中，数据稽查系统对数据提供者上传的训练数据进行有效性认证。

在步骤 302 中，数据稽查系统拒绝将有效性认证失败的训练数据存入云数据存储平台。

在步骤 303 中，检索数据平台根据数据提供者提供的训练数据，建立数据索引表。

在步骤 304 中，检索数据平台接收检索指令，根据检索指令在数据索引表中进行数据检索，并生成检索结果。

在步骤 305 中，检索数据平台接收用户终端的数据选取指令，根据数据选取指令向鉴权中心发起鉴权许可请求。

其中，鉴权许可请求包括训练数据的数据标识。

在步骤 306 中，鉴权中心接收鉴权许可请求，根据鉴权许可请求创建数据标识的数据令牌，并将数据令牌下发给云数据存储平台中的权限网关和用户终端。

在步骤 307 中，云数据存储平台中的权限网关根据下发得到的数据令牌，建立第一对应关系。

其中，第一对应关系为数据标识与数据令牌的对应关系。

在步骤 308 中，云模型训练平台接收模型训练创建指令，获取待训练模型。

在步骤 309 中，云模型训练平台生成并向云数据存储平台中的权限网关发送训练数据调用请求，以调用云数据存储平台中存储的训练数据。

其中，训练数据调用请求包括鉴权中心下发至用户终端的数据令牌。

在步骤 310 中，云数据存储平台中的权限网关接收训练数据调用请求，根据训练数据调用请求中的数据令牌，在第一对应关系中查找目标数据标识，并将目标数据标识对应的训练数据导出至云模型训练平台。

其中，目标数据标识为与训练数据调用请求中的数据令牌对应的数据标识。

在步骤 311 中，云模型训练平台利用从云数据存储平台导出的训练数据，训练待训练模型，得到训练成果模型。

在步骤 312 中，云模型存储平台保存训练成果模型。

在步骤 313 中，云模型训练平台销毁云模型训练平台内训练训练成果模型所利用的训练数据和待训练模型。

在步骤 314 中，模型推理平台接收推理请求，推理请求包括待处理数据。

在步骤 315 中，模型推理平台从镜像平台加载模型推理运行环境，并从云模型存储平台调用训练成果模型，将待处理数据导入训练成果模型进行模型推理。

图 7 为本发明一实施例中一种模型训练方法的另一种具体实现方式的流程图。图 7 与图 6 的不同之处在于，图 6 中的步骤 S307 可替换为图 7 中的步骤 S316；图 6 中的步骤 S310 可替换为图 7 中的步骤 S317 和步骤 S318。

在步骤 S316 中，云数据存储平台中的权限网关根据下发得到的数据令牌，建立第二对应关系。

其中，第二对应关系为数据令牌与数据路由的对应关系。数据路由包括训练数据的统一资源定位符路径。

在步骤 S317 中，云数据存储平台中的权限网关接收训练数据调用请求，根据训练数据调用请求中的数据令牌，在第二对应关系中查找目标数据路由。

其中，目标数据路由为与训练数据调用请求中的数据令牌对应的数据路由。

在步骤 S318 中，云数据存储平台中的权限网关访问目标数据存储服务器，以将目标数据存储服务器中目标数据路由指示的训练数据导出至云模型训练平台。

其中，目标数据存储服务器为与目标数据路由对应的数据存储服务器。

在一个示例中，还可以根据具体场景对数据令牌进行更新，从而保证训练数据的安全。权限网关获取更新判断参数，判断更新判断参数是否满足更新条件。若判定更新判断参数满足更新条件，权限网关向鉴权中心发送更新请求。鉴权中心接收更新请求，根据更新请求更新数据令牌。权限网关与鉴权中心同步更新数据令牌。

示例性的，更新判断参数包括对鉴权许可请求的拒绝次数。数据令牌更新过程可具体为：权限网关监测鉴权中心对鉴权许可请求的处理过程，并获取鉴权中心对鉴权许可请求的拒绝次数，并判断鉴权中心对鉴权许可请求的拒绝次数是否超出更新条件中的拒绝次数更新阈值；若监测到鉴权中心对鉴权许可请求的拒绝次数超出更新条件中的拒绝次数更新阈值，则向鉴权中心发送更新请求。

示例性的，更新判断参数包括训练数据的调用次数。数据令牌更新过程可具体为：权限网关获取一段时长内的训练数据的调用次数，判断在一段时长内，同一训练数据的调用次数是否超出更新条件中的调用次数更新阈值；若在一段时长内，同一训练数据的调用次数超出更新条件中的调用次数更新阈值，则向鉴权中心发送更新请求。

上述方法实施例中各步骤的说明内容可参照上述系统实施例中的相关说明。

本发明实施例还可提供一种存储介质，该存储介质上存储有程序，程序被处理器执行时实现上述实施例中的模型训练方法。

权 利 要 求 书

1、一种模型训练系统，其特征在于，包括云数据存储平台和云模型训练平台；

所述云数据存储平台用于存储训练数据，以及用于接收训练数据调用请求，根据所述训练数据调用请求，将与所述数据调用指令对应的训练数据导出至所述云模型训练平台；

所述云模型训练平台用于接收模型训练创建指令，获取待训练模型，以及用于生成并向所述云数据存储平台发送训练数据调用请求，以及用于利用从所述云数据存储平台导出的训练数据，训练所述待训练模型，得到训练成果模型。

2、根据权利要求 1 所述的系统，其特征在于，还包括检索数据平台和鉴权中心；所述云数据存储平台包括权限网关；

所述检索数据平台用于根据数据提供者提供的训练数据，建立数据索引表，以及用于接收检索指令，根据所述检索指令在所述数据索引表中进行数据检索，并生成检索结果，以及用于接收用户终端针对所述检索结果的数据选取指令，根据所述数据选取指令向所述鉴权中心发起鉴权许可请求，所述鉴权许可请求包括所述训练数据的数据标识；

所述鉴权中心用于接收所述鉴权许可请求，根据所述鉴权许可请求创建所述数据标识的数据令牌，并将所述数据令牌下发给所述权限网关和所述用户终端；

所述云模型训练平台还用于向所述权限网关发送所述训练数据调用请求，所述训练数据调用请求包括所述鉴权中心下发至所述用户终端的数据令牌；

所述权限网关用于建立第一对应关系，所述第一对应关系为所述数据标识与所述数据令牌一一对应的关系，以及用于接收所述训练数据调用请求，根据所述训练数据调用请求中的数据令牌，在所述第一对应关系中查找目标数据标识，所述目标数据标识为与所述训练数据调用请求中的数据令牌对应的数据标识，以及用于将所述目标数据标识对应的训练数据导出至所述云模型训练平台。

3、根据权利要求 1 所述的系统，其特征在于，还包括检索数据平台和鉴权中心；所述云数据存储平台包括权限网关和至少一个数据存储服务器；

所述检索数据平台用于根据数据提供者提供的训练数据，建立数据索引表，以及接收检索指令，根据所述检索指令在所述数据索引表中进行数据检索，并生成检索结果，以及用于接收用户终端针对所述检索结果的数据选取指令，根据所述数据选取指令向所述鉴权中心发起鉴权许可请求，所述鉴权许可请求包括所述训练数据的数据标识；

所述鉴权中心用于接收所述鉴权许可请求，根据所述鉴权许可请求创建所述数据标识的数据令牌，并将所述数据令牌下发给所述权限网关和所述用户终端；

所述云模型训练平台还用于向所述权限网关发送所述训练数据调用请求，所述训练数据调用请求包括所述鉴权中心下发至所述用户终端的数据令牌；

所述权限网关用于建立第二对应关系，所述第二对应关系为所述数据令牌与数据路由的对应关系，所述数据路由包括所述训练数据的统一资源定位符路径，以及用于接收所述训练数据调用请求，根据所述训练数据调用请求中的数据令牌，在所述第二对应关系中查找目标数据路由，所述目标数据路由为与所述训练数据调用请求中的数据令牌对应的数据路由，以及用于访问目标数据存储服务器，以将所述目标数据存储服务器中所述目标数据路由指示的训练数据导出至所述云模型训练平台，所述目标数据存储服务器为与所述目标

数据路由对应的数据存储服务器。

4、根据权利要求 3 所述的系统，其特征在于，还包括访问路由器，所述权限网关通过所述访问路由器中预定的标准访问接口从所述目标数据存储服务器中导出所述目标数据路由指示的训练数据。

5、根据权利要求 2 或 3 所述的系统，其特征在于，

所述权限网关还用于获取更新判断参数，判断所述更新判断参数是否满足更新条件，以及用于若判定所述更新判断参数满足更新条件，向所述鉴权中心发送更新请求，以及用于与所述鉴权中心同步更新数据令牌；

所述鉴权中心还用于接收所述更新请求，根据所述更新请求更新所述数据令牌。

6、根据权利要求 5 所述的系统，其特征在于，所述更新判断参数包括对所述鉴权许可请求的拒绝次数；

所述权限网关还用于监测所述鉴权中心对所述鉴权许可请求的处理过程，以及用于若监测到所述鉴权中心对所述鉴权许可请求的拒绝次数超出所述更新条件中的拒绝次数更新阈值，则向所述鉴权中心发送所述更新请求。

7、根据权利要求 6 所述的系统，其特征在于，所述更新判断参数包括训练数据的调用次数；

所述权限网关还用于获取一段时长内的所述训练数据的调用次数，以及用于若在所述一段时长内，同一所述训练数据的调用次数超出所述更新条件中的调用次数更新阈值，则向所述鉴权中心发送所述更新请求。

8、根据权利要求 1 所述的系统，其特征在于，所述云模型训练平台还用于训练得到所述训练成果模型后，销毁所述云模型训练平台内训练所述训练成果模型所利用的训练数据和待训练模型。

9、根据权利要求 1 所述的系统，其特征在于，还包括数据稽查系统；

所述数据稽查系统用于对数据提供者上传的训练数据进行有效性认证，拒绝将有效性认证失败的训练数据存入所述云数据存储平台。

10、根据权利要求 1 所述的系统，其特征在于，还包括云模型存储平台；

所述云模型存储平台用于提供所述待训练模型，以及保存所述训练成果模型。

11、根据权利要求 10 所述的系统，其特征在于，还包括镜像平台和模型推理平台；

所述镜像平台用于存储模型推理运行环境；

所述模型推理平台用于接收推理请求，所述推理请求包括待处理数据，以及从所述镜像平台加载所述模型推理运行环境，以及从所述云模型存储平台调用所述训练成果模型，将所述待处理数据导入所述训练成果模型进行模型推理。

12、一种模型训练方法，其特征在于，包括：

云模型训练平台接收模型训练创建指令，获取待训练模型；

所述云模型训练平台生成并向云数据存储平台发出训练数据调用请求，以调用所述云数据存储平台中存储的训练数据；

所述云数据存储平台接收所述训练数据调用请求，将与所述训练数据调用请求对应的训练数据导出至所述云模型训练平台；

所述云模型训练平台利用从云数据存储平台导出的训练数据，训练所述待训练模型，

得到训练成果模型。

13、根据权利要求 12 所述的方法，其特征在于，还包括:

检索数据平台根据数据提供者提供的训练数据，建立数据索引表;

所述检索数据平台接收检索指令，根据所述检索指令在所述数据索引表中进行数据检索，并生成检索结果;

所述检索数据平台接收用户终端的数据选取指令，根据所述数据选取指令向鉴权中心发起鉴权许可请求，所述鉴权许可请求包括所述训练数据的数据标识;

所述鉴权中心接收所述鉴权许可请求，根据所述鉴权许可请求创建所述数据标识的数据令牌，并将所述数据令牌下发给权限网关和所述用户终端;

所述权限网关根据下发得到的所述数据令牌，建立第一对应关系，所述第一对应关系为所述数据标识与所述数据令牌一一对应的关系。

14、根据权利要求 13 所述的方法，其特征在于，所述云模型训练平台生成并向所述云数据存储平台发送训练数据调用请求，包括:

所述云模型训练平台生成并向所述权限网关发送所述训练数据调用请求，所述训练数据调用请求包括所述鉴权中心下发至所述用户终端的数据令牌;

所述云数据存储平台接收所述训练数据调用请求，将与所述训练数据调用请求对应的训练数据导出至所述云模型训练平台，包括:

所述云数据存储平台中的所述权限网关接收所述训练数据调用请求，根据所述训练数据调用请求中的数据令牌，在所述第一对应关系中查找目标数据标识，并将所述目标数据标识对应的训练数据导出至所述云模型训练平台，所述目标数据标识为与所述训练数据调用请求中的数据令牌对应的数据标识。

15、根据权利要求 12 所述的方法，其特征在于，还包括:

检索数据平台根据数据提供者提供的训练数据，建立数据索引表;

所述检索数据平台接收检索指令，根据所述检索指令在所述数据索引表中进行数据检索，并生成并发送检索结果;

所述检索数据平台接收用户终端针对所述检索结果的数据选取指令，根据所述数据选取指令向鉴权中心发起鉴权许可请求，所述鉴权许可请求包括所述训练数据的数据标识;

所述鉴权中心接收所述鉴权许可请求，根据所述鉴权许可请求创建所述数据标识的数据令牌，并将所述数据令牌下发给权限网关和所述用户终端;

所述权限网关根据下发得到的所述数据令牌，建立第二对应关系，所述第二对应关系为所述数据令牌与数据路由的对应关系，所述数据路由包括所述训练数据的统一资源定位符路径。

16、根据权利要求 15 所述的方法，其特征在于，所述云模型训练平台生成并向所述云数据存储平台发送训练数据调用请求，包括:

所述云模型训练平台生成并向所述权限网关发送所述训练数据调用请求，所述训练数据调用请求包括所述鉴权中心下发至所述用户终端的数据令牌;

所述云数据存储平台接收所述训练数据调用请求，将与所述训练数据调用请求对应的训练数据导出至所述云模型训练平台，包括:

所述云数据存储平台中的所述权限网关接收所述训练数据调用请求，根据所述训练数

据调用请求中的数据令牌，在所述第二对应关系中查找目标数据路由，所述目标数据路由为与所述训练数据调用请求中的数据令牌对应的数据路由；

所述权限网关访问目标数据存储服务器，以将所述目标数据存储服务器中所述目标数据路由指示的训练数据导出至所述云模型训练平台，所述目标数据存储服务器为与所述目标数据路由对应的数据存储服务器。

17、根据权利要求 12 至 16 中任意一项所述的方法，其特征在于，还包括：

所述权限网关获取更新判断参数，判断所述更新判断参数是否满足更新条件；

若判定所述更新判断参数满足更新条件，所述权限网关向所述鉴权中心发送更新请求；

所述鉴权中心接收所述更新请求，根据所述更新请求更新所述数据令牌；

所述权限网关与所述鉴权中心同步更新数据令牌。

18、根据权利要求 17 所述的方法，其特征在于，所述更新判断参数包括对所述鉴权许可请求的拒绝次数；

所述权限网关获取更新判断参数，判断所述更新判断参数是否满足更新条件，包括：

所述权限网关监测所述鉴权中心对所述鉴权许可请求的处理过程，并获取所述鉴权中心对所述鉴权许可请求的拒绝次数，并判断所述鉴权中心对所述鉴权许可请求的拒绝次数是否超出所述更新条件中的拒绝次数更新阈值；

所述若判定所述更新判断参数满足更新条件，所述权限网关向所述鉴权中心发送更新请求，包括：

若监测到所述鉴权中心对所述鉴权许可请求的拒绝次数超出所述更新条件中的拒绝次数更新阈值，则向所述鉴权中心发送所述更新请求。

19、根据权利要求 17 所述的方法，其特征在于，所述更新判断参数包括训练数据的调用次数；

所述权限网关获取更新判断参数，判断所述更新判断参数是否满足更新条件，包括：

所述权限网关获取一段时长内的所述训练数据的调用次数，判断在所述一段时长内，同一所述训练数据的调用次数是否超出所述更新条件中的调用次数更新阈值；

所述若判定所述更新判断参数满足更新条件，所述权限网关向所述鉴权中心发送更新请求，包括：

若在所述一段时长内，同一所述训练数据的调用次数超出所述更新条件中的调用次数更新阈值，则向所述鉴权中心发送所述更新请求。

20、根据权利要求 12 所述的方法，其特征在于，在所述云模型训练平台利用从云数据存储平台导出的训练数据，训练所述待训练模型，得到训练成果模型之后，还包括：

所述云模型训练平台销毁所述云模型训练平台内训练所述训练成果模型所利用的训练数据和待训练模型。

21、根据权利要求 12 所述的方法，其特征在于，还包括：

数据稽查系统对数据提供者上传的训练数据进行有效性认证；

所述数据稽查系统拒绝将有效性认证失败的训练数据存入所述云数据存储平台。

22、根据权利要求 12 所述的方法，其特征在于，在所述云模型训练平台利用从云数据存储平台导出的训练数据，训练所述待训练模型，得到训练成果模型之后，还包括：

所述云模型存储平台保存所述训练成果模型。

23、根据权利要求 22 所述的方法，其特征在于，还包括:

模型推理平台接收推理请求，所述推理请求包括待处理数据;

所述模型推理平台从所述镜像平台加载所述模型推理运行环境，并从所述云模型存储平台调用所述训练成果模型，将所述待处理数据导入所述训练成果模型进行模型推理。

24、一种存储介质，其特征在于，所述存储介质上存储有程序，所述程序被处理器执行时实现如权利要求 12 至 23 中任意一项所述的模型训练方法。

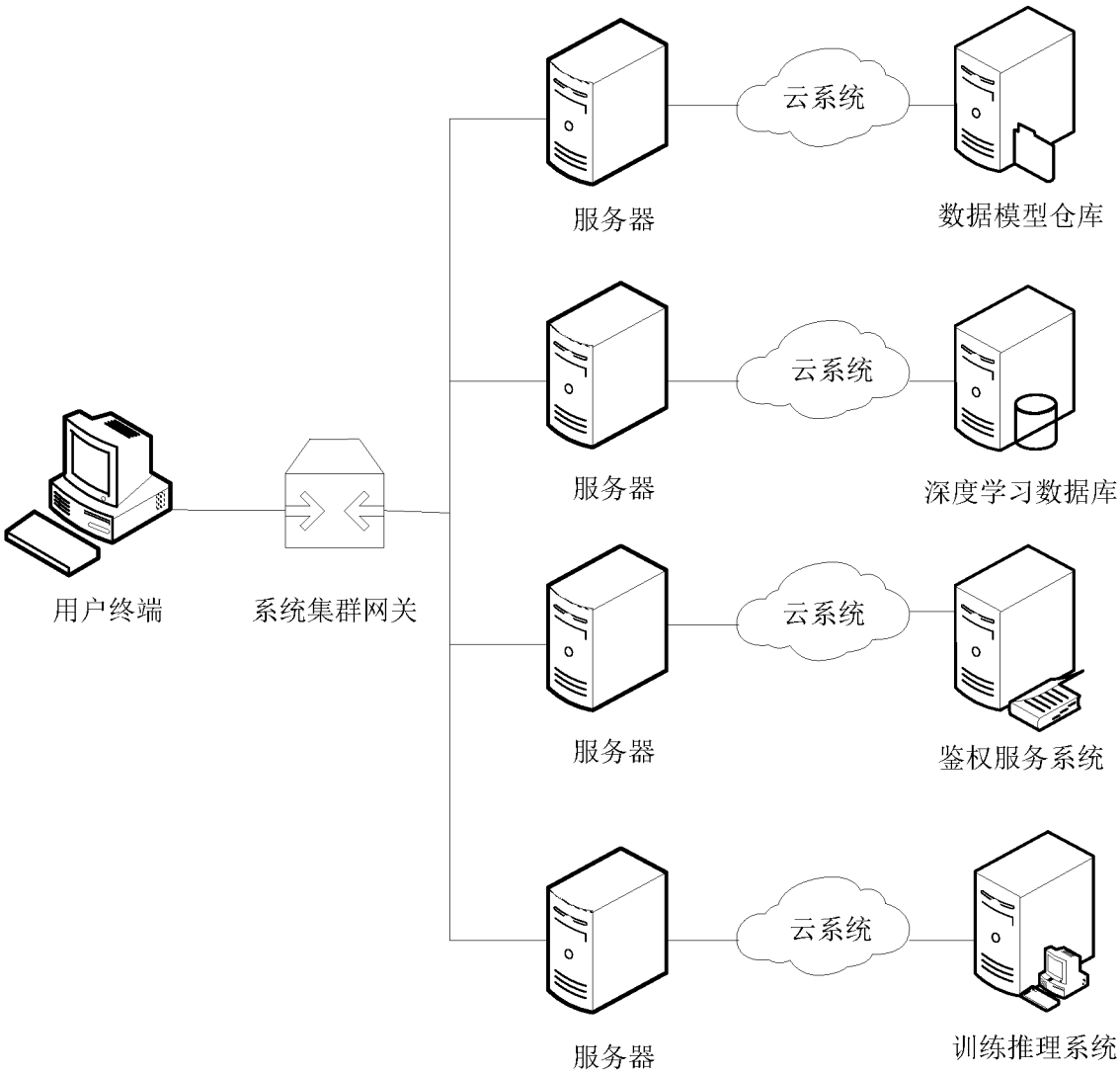
云系统
服务器
数据模型仓库
云系统
服务器
深度学习数据库
用户终端
系统集群网关
云系统
服务器
鉴权服务系统
云系统
服务器
训练推理系统

图 1


10
用户终端20
云数据存储平台11
云模型训练平台12

图 2

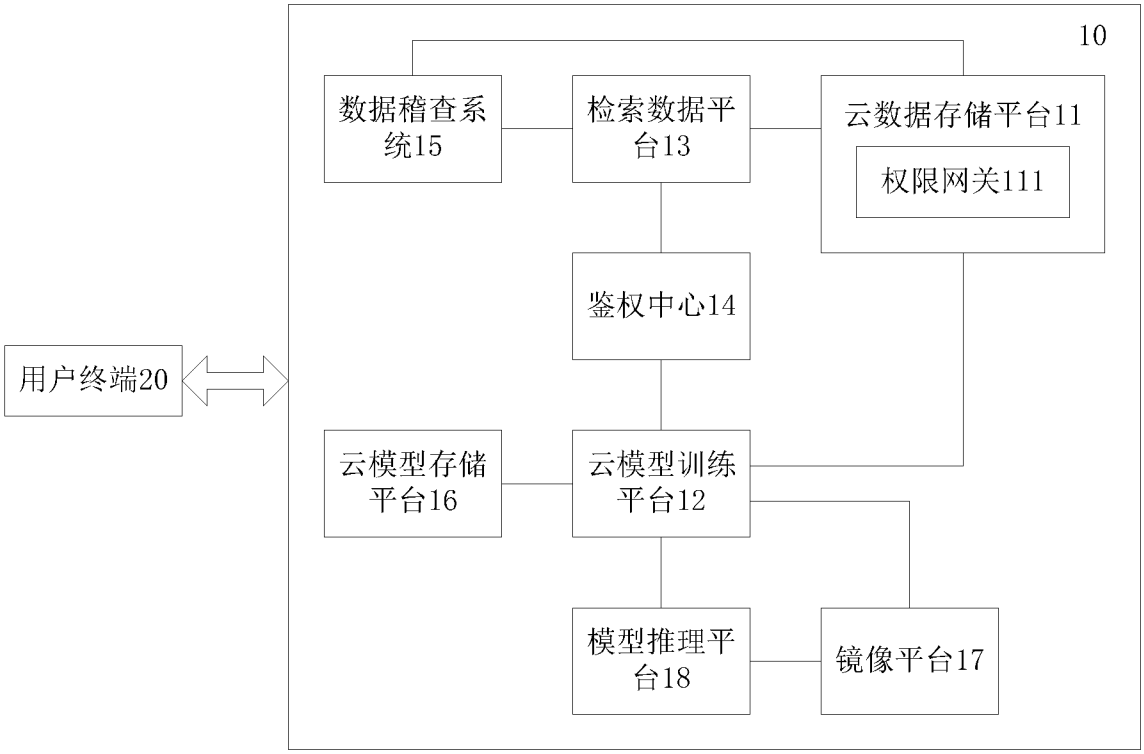
10
数据稽查系统15
检索数据平台13
云数据存储平台11
权限网关111
鉴权中心14
用户终端20
云模型存储平台16
云模型训练平台12
模型推理平台18
镜像平台17

图 3

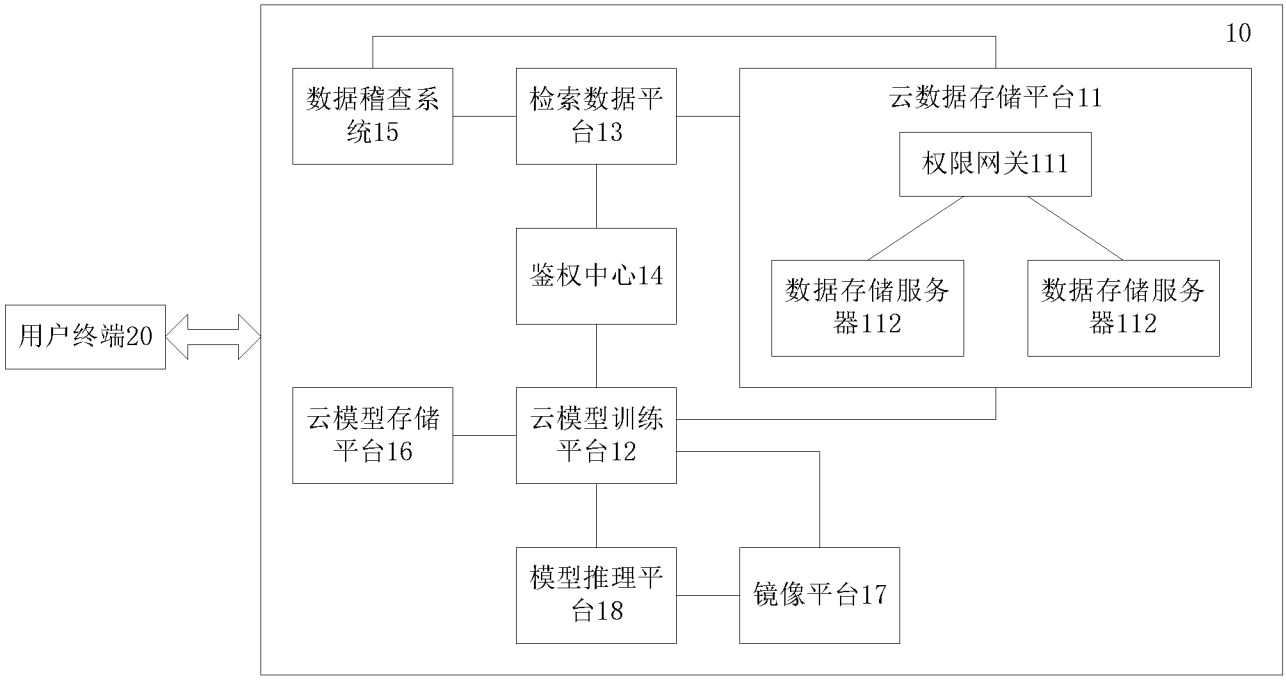
10
数据稽查系统15
检索数据平台13
云数据存储平台11
权限网关111
数据存储服务器112
数据存储服务器112
鉴权中心14
用户终端20
云模型存储平台16
云模型训练平台12
模型推理平台18
镜像平台17

图 4

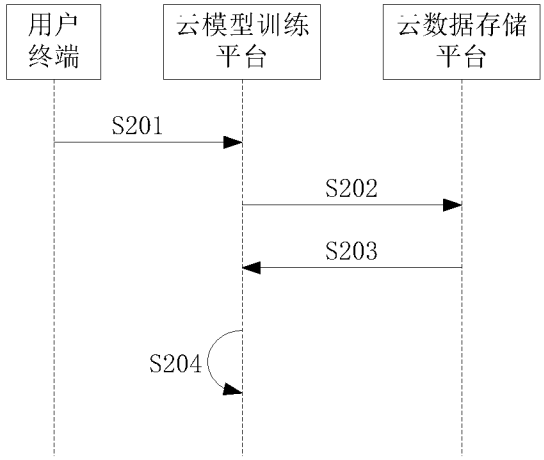
用户终端
云模型训练平台
云数据存储平台
S201
S202
S203
S204

图 5

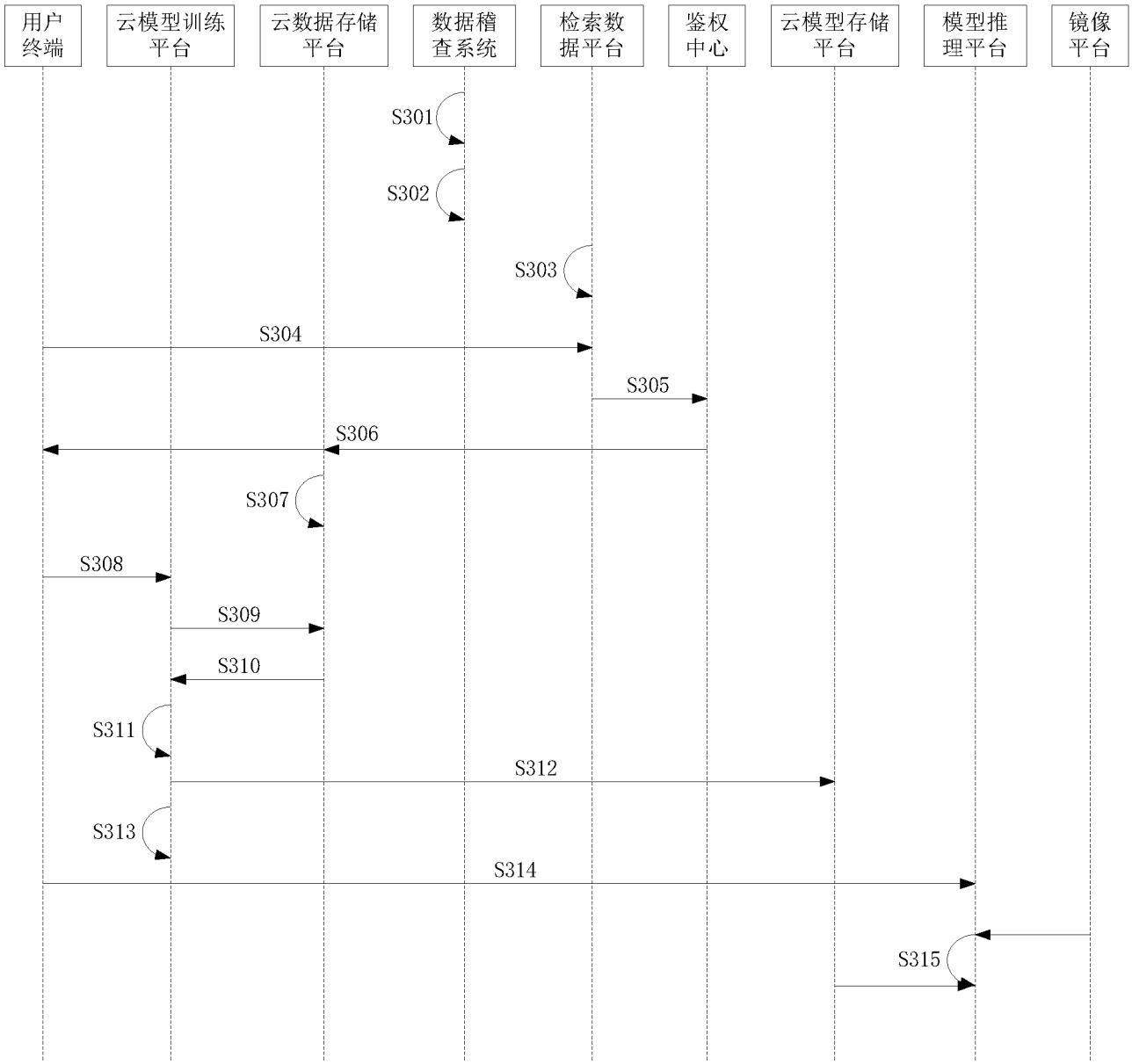
用户终端
云模型训练平台
云数据存储平台
数据稽查系统
检索数据平台
鉴权中心
云模型存储平台
模型推理平台
镜像平台
S301
S302
S303
S304
S305
S306
S307
S308
S309
S310
S311
S312
S313
S314
S315

图 6

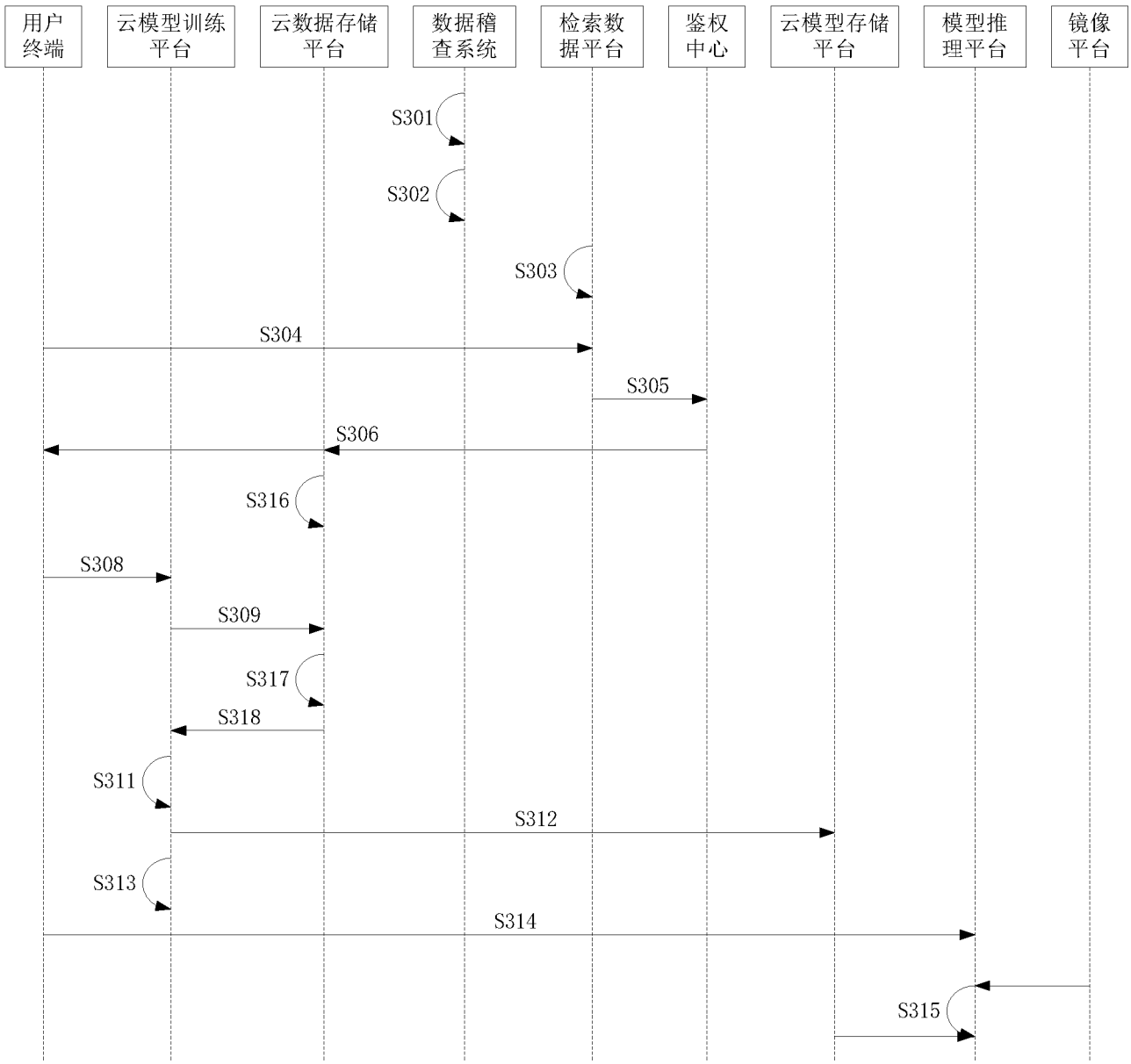
用户终端
云模型训练平台
云数据存储平台
数据稽查系统
检索数据平台
鉴权中心
云模型存储平台
模型推理平台
镜像平台
S301
S302
S303
S304
S305
S306
S316
S308
S309
S317
S318
S311
S312
S313
S314
S315

图 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/114082

A. CLASSIFICATION OF SUBJECT MATTER

G06N 3/08(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06N3/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

USTXT; EPTXT; CNTXT; CNABS; WOTXT; VEN; CNKI: 云, 模型, 训练, 存储, 服务器, 平台, 索引, 检索, 鉴权, 更新, cloud, model, train, storage, server, platform, index, search, authentication, update

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 106204780 A (WUHAN UNIVERSITY OF TECHNOLOGY) 07 December 2016 (2016-12-07) description, paragraphs 31, 39, and 52	1, 8-12, 20-24
Y	CN 106204780 A (WUHAN UNIVERSITY OF TECHNOLOGY) 07 December 2016 (2016-12-07) description, paragraphs 31, 39, and 52	2-7, 13-19
Y	CN 105575389 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 11 May 2016 (2016-05-11) description, paragraphs 8-12	2-7, 13-19
A	CN 103389719 A (LINYI TOP NETWORK CO., LTD.) 13 November 2013 (2013-11-13) entire document	1-24
A	CN 107195186 A (QIANXUN SPATIAL INTELLIGENCE INC.) 22 September 2017 (2017-09-22) entire document	1-24
A	US 2017154113 A1 (WAL-MART STORES, INC.) 01 June 2017 (2017-06-01) entire document	1-24

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 January 2019	**31 January 2019**

Name and mailing address of the ISA/CN	Authorized officer
National Intellectual Property Administration, PRC (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/114082

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106204780	A	07 December 2016	None			
CN	105575389	A	11 May 2016	None			
CN	103389719	A	13 November 2013	CN	103389719	B	22 July 2015
CN	107195186	A	22 September 2017	None			
US	2017154113	A1	01 June 2017	None			

国际检索报告

国际申请号
PCT/CN2018/114082

A. 主题的分类

G06N 3/08(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06N3/-

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用）)

USTXT;EPTXT;CNTXT;CNABS;WOTXT;VEN;CNKI:云，模型，训练，存储，服务器，平台，索引，检索，鉴权，更新，cloud, model, train, storage, server, platform, index, search, authentication, update

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
X	CN 106204780 A (武汉理工大学) 2016年 12月 7日 (2016-12-07) 说明书第31、39、52段	1、8-12、20-24
Y	CN 106204780 A (武汉理工大学) 2016年 12月 7日 (2016-12-07) 说明书第31、39、52段	2-7、13-19
Y	CN 105575389 A (百度在线网络技术北京有限公司) 2016年 5月 11日 (2016-05-11) 说明书第8-12段	2-7、13-19
A	CN 103389719 A (临沂市拓普网络股份有限公司) 2013年 11月 13日 (2013-11-13) 全文	1-24
A	CN 107195186 A (千寻位置网络有限公司) 2017年 9月 22日 (2017-09-22) 全文	1-24
A	US 2017154113 A1 (WAL-MART STORES, INC.) 2017年 6月 1日 (2017-06-01) 全文	1-24

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

"A" 认为不特别相关的表示了现有技术一般状态的文件

"E" 在国际申请日的当天或之后公布的在先申请或专利

"L" 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

"O" 涉及口头公开、使用、展览或其他方式公开的文件

"P" 公布日先于国际申请日但迟于所要求的优先权日的文件

"T" 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

"X" 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

"Y" 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

"&" 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2019年 1月 7日	2019年 1月 31日

ISA/CN的名称和邮寄地址	受权官员
中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088	夏雪
传真号 (86-10)62019451	电话号码 86-(20)-28950718

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/114082

检索报告引用的专利文件			公布日（年/月/日）	同族专利			公布日（年/月/日）
CN	106204780	A	2016年 12月 7日	无			
CN	105575389	A	2016年 5月 11日	无			
CN	103389719	A	2013年 11月 13日	CN	103389719	B	2015年 7月 22日
CN	107195186	A	2017年 9月 22日	无			
US	2017154113	A1	2017年 6月 1日	无			