

说明书

一种利用模型生成内容的方法、系统及集群

本申请要求于 2023 年 6 月 28 日提交中国国家知识产权局、申请号为 202310777021.3、申请名称为“一种命令传输方法”的中国专利申请的优先权，以及于 2023 年 9 月 20 日提交中国国家知识产权局、申请号为 202311218468.3、申请名称为“一种利用模型生成内容的方法、系统及集群”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及云计算（cloud computing）和人工智能（artificial intelligence, AI）技术领域，特别涉及一种利用模型生成内容的方法、系统及集群。

背景技术

人工智能生成内容（AI-generated content, AIGC）是一种通过AI模型来生成数字内容（例如图片、音频和文本）的技术。其中，用于生成数字内容的AI模型可称为生成式（generative model, GM）模型。随着生成式模型的发展，普通用户通过AIGC，可生成高质量的内容。因此，AIGC得到了迅速应用。

如图1所示，由于生成式模型对算力要求较高，通常将生成式模型部署在云端。由此，用户使用轻量级终端设备，就可以使用生成式模型生成内容。具体而言，用户使用终端设备，向云端的生成式模型发送内容生成指令。生成式模型根据该内容生成指令，生成用户可能想要的内容，即目标内容。

内容生成指令可能包含用户的敏感信息，例如，用户可能无意间将个人信息、商业秘密或者反映用户喜好的信息等敏感信息包含在了内容生成指令中。因此，向云端发送内容生成指令可能使得云端获取用户的敏感信息，导致用户的敏感信息泄露。

发明内容

本申请实施例提出了一种利用模型生成内容的方法、系统及集群，可以避免云端获取内容生成指令中的用户敏感信息，保护用户的信息安全。

第一方面，提供了一种利用模型生成内容的方法，应用于内容生成系统，系统包括客户端和服务端，模型包括第一模块和第二模块，客户端和第一模块部署在用户本地，服务端和第二模块部署在云端，并且，第一模块的输出为第二模块的输入；该方法包括：客户端接收用户发出的内容生成指令，内容生成指令用于指示模型按照用户的意图生成目标内容；客户端将内容生成指令输入至第一模块，使得第一模块输出模型生成目标内容的中间结果；服务端接收客户端发送的中间结果，并将中间结果输入至第二模块，使得第二模块输出目标内容；服务端将目标内容发送至客户端。

该方法将用于生成目标内容的模型一部分部署在用户本地，另一部分部署在云端。部署在用户本地的客户端在接收到内容生成指令时，可以利用本地部署的部分模型，对内容生成指令进行处理，得到难以用于倒推出内容生成指令的中间结果。然后，将中间结果发送至云端，使得云端基于中间结果生成目标内容，从而在无需用户本地运行全部模型的同时，避免了内容生成指令中的敏感信息泄露。

在一种可能的实现方式中，模型为包括多个网络层的生成式模型，多个网络层在生成式模型的数据传递方向上依次排列，其中，第一模块包括多个网络层中的至少一个网络层，第二模块包括多个网络层中除至少一个网络层之外的网络层。

在该实现方式中，可以将生成式模型的一部分网络层部署在用户本地，另一部分网络层部署在云端。部署在用户本地的网络层可以对用户发出的内容生成指令进行处理，得到生成式模型的中间结果。由于AI模型的不可解释性，中间结果难以倒推出模型的输入，即内容生成指令，因此，云端在获取到的内容生成指令后，也难以获取用户的内容生成指令，进而难以获取内容生成指令可能包含的敏感信息。

在一种可能的实现方式中，内容生成指令包括用户的意图，模型包括解释器和生成器，解释器用于提取用户的意图的特征，得到编码向量，生成器用于基于编码向量，生成目标内容；其中，第一模块包括解释器，第二模块包括生成器；或者，第一模块包括解释器中的第一网络层，第二模块包括解释器中的第二网络层和生成器；第一网络层的输出为第二网络层的输入；或者，第一模块包括解释器和生成器中的第三网络层，第二模块包括生成器中的第四网络层；第三网络层的输出为第四网络层的输入。

在该实现方式中，第一模块包括解释器、部分解释器、或者全部的解释器以及部分生成器，也就是说，可以灵活配置第一模块。其中，当用户本地的算力较高时，可以使得第一模块包括全部的解释器，

或者全部的解释器以及部分生成器，从而可以充分利用用户本地的算力，并且，第一模块输出的中间结果也更加难以用于推测内容生成指令。当用户本地的算力较低时，可以使得第一模块包括部分的解释器，从而降低对本地算力的需求。

在一种可能的实现方式中，模型为生成式模型，该方法还包括：服务端训练模型；在模型训练完成后，服务端将第一模块发送至用户本地。

在该实现方式中，模型可以在云端训练，在模型训练完成后，将模型中的第一模块发送至用户本地。由此，无需用户训练或设计第一模块，提供了用户体验。

在一种可能的实现方式中，该方法还包括：客户端接收用户发出的参数调整操作，并基于参数调整操作，调整第一模块的参数。

在该实现方式中，用户可以调整或者说修改第一模块的参数，使得云端或其他方式不感知第一模块的参数，从而无法利用第一模块的参数，来推测用户发出的内容生成指令。

在一种可能的实现方式中，第一模块包括用户设置的敏感信息过滤器，过滤器用于拦截内容生成指令中的敏感信息，得到中间结果。

在该实现方式中，第一模块可实现为过滤器，过滤器对算力要求低，从而使得用户使用轻量级的用户本地设备即可通过本申请实施例提供的方法生成内容。

第二方面，提供了一种内容生成系统，系统用于利用模型生成内容，其中，系统包括客户端和服务端，模型包括第一模块和第二模块，客户端和第一模块部署在用户本地，服务端和第二模块部署在云端，并且，第一模块的输出为第二模块的输入；客户端用于接收用户发出的内容生成指令，内容生成指令用于指示模型按照用户的意图生成目标内容；客户端用于将内容生成指令输入至第一模块，使得第一模块输出模型生成目标内容的中间结果；服务端用于接收客户端发送的中间结果，并将中间结果输入至第二模块，使得第二模块输出目标内容；服务端用于将目标内容发送至客户端。

在一种可能的实现方式中，模型为包括多个网络层的生成式模型，多个网络层在生成式模型的数据传递方向上依次排列，其中，第一模块包括多个网络层中的至少一个网络层，第二模块包括多个网络层中除至少一个网络层之外的网络层。

在一种可能的实现方式中，内容生成指令包括用户的意图，模型包括解释器和生成器，解释器用于提取用户的意图的特征，得到编码向量，生成器用于基于编码向量，生成目标内容；其中，第一模块包括解释器，第二模块包括生成器；或者，第一模块包括解释器中的第一网络层，第二模块包括解释器中的第二网络层和生成器；第一网络层的输出为第二网络层的输入；或者，第一模块包括解释器和生成器中的第三网络层，第二模块包括生成器中的第四网络层；第三网络层的输出为第四网络层的输入。

在一种可能的实现方式中，模型为生成式模型，服务端还用于：训练模型；在模型训练完成后，将第一模块发送至用户本地。

在一种可能的实现方式中，客户端还用于：接收用户发出的参数调整操作，并基于参数调整操作，调整第一模块的参数。

在一种可能的实现方式中，第一模块包括用户设置的敏感信息过滤器，过滤器用于拦截内容生成指令中的敏感信息，得到中间结果。

第三方面，提供了一种利用模型生成内容的方法，应用于部署在用户本地的客户端，客户端所在的内容生成系统还包括部署在云端的服务端，模型包括第一模块和第二模块，第一模块部署在用户本地，第二模块部署在云端，并且，第一模块的输出为第二模块的输入；该方法包括：客户端接收用户发出的内容生成指令，内容生成指令用于指示模型按照用户的意图生成目标内容；客户端将内容生成指令输入至第一模块，使得第一模块输出模型生成目标内容的中间结果；客户端将中间结果发送至服务端；其中，服务端用于将中间结果输入至第二模块，使得第二模块输出目标内容；客户端接收服务端发送的目标内容。

在一种可能的实现方式中，该方法还包括：客户端接收用户发出的参数调整操作，并基于参数调整操作，调整第一模块的参数。

第四方面，提供了一种利用模型生成内容的客户端，客户端部署在用户本地，客户端所在的内容生成系统还包括部署在云端的服务端，模型包括第一模块和第二模块，第一模块部署在用户本地，第二模块部署在云端，并且，第一模块的输出为第二模块的输入；客户端包括：接收模块，用于接收用户发出的内容生成指令，内容生成指令用于指示模型按照用户的意图生成目标内容；输入模块，用于将内容生

成指令输入至第一模块，使得第一模块输出模型生成目标内容的中间结果；通信模块，用于将中间结果发送至服务端；其中，服务端用于将中间结果输入至第二模块，使得第二模块输出目标内容；通信模块还用于接收服务端发送的目标内容。

在一种可能的实现方式中，接收模块还用于：接收用户发出的参数调整操作，并基于参数调整操作，调整第一模块的参数。

第五方面，提供了一种计算设备集群，包括至少一个计算设备，每个计算设备包括处理器和存储器；至少一个计算设备的处理器用于执行至少一个计算设备的存储器中存储的指令，以使得计算设备集群执行第三方面提供的方法。

第六方面，提供了一种计算机可读存储介质，包括计算机程序指令，当计算机程序指令由计算设备集群执行时，计算设备集群执行第三方面提供的方法。

第七方面，提供了一种包含指令的计算机程序产品，当指令被计算机设备集群运行时，使得计算机设备集群执行第三方面提供的方法。

第二至第七方面的有益效果可以参考上文对第一方面的有益效果的介绍，在此不再赘述。

附图说明

- 图 1 为一种内容生成方案的示意图；
- 图 2 为一种生成式模型的结构示意图；
- 图 3A 为本申请实施例提供的一种系统架构的示意图；
- 图 3B 为本申请实施例提供的一种系统架构的示意图；
- 图 3C 为本申请实施例提供的一种系统架构的示意图；
- 图 4 为本申请实施例提供的一种内容生成方法的流程图；
- 图 5 为本申请实施例提供的一种模块 210 和模块 220 的示意图；
- 图 6 为本申请实施例提供的一种模块 210 和模块 220 的示意图；
- 图 7 为本申请实施例提供的一种模块 210 和模块 220 的示意图；
- 图 8 为本申请实施例提供的一种模块 210 和模块 220 的示意图；
- 图 9 为本申请实施例提供的一种内容生成系统的结构示意图；
- 图 10 为本申请实施例提供的一种内容生成方法的流程图；
- 图 11 为本申请实施例提供的一种客户端的结构示意图；
- 图 12 为本申请实施例提供的一种计算设备的结构示意图；
- 图 13 为本申请实施例提供的一种计算设备集群的结构示意图；
- 图 14 为本申请实施例提供的一种计算设备集群的结构示意图。

具体实施方式

下面将结合附图对本申请实施例提供的方案进行描述。其中，在本申请实施例中，“多个”是指两个或两个以上。

为便于理解本申请实施例提供的方案，在展开介绍本申请实施例提供的方案之前，先介绍本申请实施例可能涉及到的一些技术用语。

内容生成指令：是指用户发出的、用于指示生成式模型按照用户意图，生成内容的信息或命令。内容生成指令包括用户意图。用户意图用于描述目标内容（即所要生成的内容），例如，用户想要生成一幅有关带烟囱的房子的画，则用户意图可以为“画一座带烟囱的房子”。

生成式模型：用于理解用户意图，并按照用户意图生成内容的神经网络。换言之，生成式模型以内容生成指令为输入，并输出符合该用户意图的内容。如图 2 所示，生成式模型通常由解释器和生成器组成。其中，解释器用于理解内容生成指令包含的用户意图，即提取用户意图的特征，得到用户意图的编码向量。用户意图的编码向量也称为用户意图的特征向量。生成器用于按照用户意图的编码向量，利用生成器自身的知识以及掌握的信息，生成符合用户意图的内容。如图 2 所示，解释器和生成器分别由多个网络层依次排列组成。其中，网络层可简称为层，每层的输出是该层的下一层的输入。每层包括多个神经元，每个神经元利用参数对数据进行计算，并输出计算结果。计算结果通常使用向量表示，可称为编码向量。其中，神经元的参数经过训练得到，模型的训练或者神经网络的训练本质是对神经元的参数的训练或者调整。目前，典型的生成式模型有文本到文本的变换器（text-to-text transfer

transformer, T5)、生成式对抗网络(generative adversarial networks, GANs)、变分自编码(variational autoencoders, VAEs)等。

云端: 相对用户本地而言的远端装置、设备或平台, 具体可以是部署在云端的一个或多个服务器或计算实例, 例如数据中心、服务器集群等。

用户本地: 是指用户侧装置、设备或平台, 具体可以是用户的终端设备, 也可以是云厂商部署在用户侧的边缘设备或计算实例。

数据脱敏(data masking): 对敏感信息(身份信息、资产信息等)通过脱敏规则进行数据的变形, 保障敏感信息不外泄, 实现对敏感信息的可靠保护。其中, 敏感信息也可以称为隐私信息。

深度学习(deep learning): 一类基于深层次神经网络算法的机器学习技术, 其主要特征是使用多重非线性变换, 构成对数据进行处理和分析。深度学习主要应用于人工智能领域的感知、决策以及内容生成等场景, 例如图像识别、语音识别、自然语言翻译、计算机博弈、内容生成等。

图 3A 示出了本申请实施例提供的一种系统架构。该系统架构包括内容生成系统 100 以及模型 200。内容生成系统 100 由客户端 110 和服务端 120 组成, 模型 200 由模块 210 和模块 220 组成。其中, 客户端 110 和模块 210 部署在用户本地, 服务端 120 和模块 220 部署在云端。

客户端 110 可以接收用户发出的内容生成指令, 并将内容生成指令发送至模块 210。模块 210 可以对内容生成指令进行数据处理, 得到中间结果。其中, 数据处理包括数据脱敏, 也就是说, 中间结果为脱敏后的数据。客户端 110 可以获取中间结果, 并通过网络, 将中间结果发送至服务端 120。服务端 120 将从客户端 110 发送的内容生成指令(即中间结果)输入至模块 220。模块 220 可以基于内容生成指令, 生成并输出目标内容。服务端 120 可获取目标内容, 并通过网络, 将目标内容发送至客户端 110。客户端 110 可以向用户展示或播放目标内容。

在本申请实施例中, 中间结果是指模型 200 生成目标内容的中间结果。具体而言, 从内容生成指令输入模型 200 到模型 200 输出目标内容的过程包括多个处理步骤, 例如, 模型包括多个网络层, 每个网络层对数据的处理为一个处理步骤。除该多个处理步骤中最后一个处理步骤外, 每个处理步骤的处理结果均可称为中间结果。

在一些实施例中, 客户端 110 或服务端 120 可以通过软件方式实现, 例如客户端 110 可以是运行在用户本地的应用(application)或进程(process), 服务端 120 可以是运行在用户本地的应用或进程。在一些实施例中, 客户端 110 或服务端 120 可以通过硬件方式实现, 例如客户端 110 可以是终端设备, 服务端 120 可以为服务器等。本申请实施例对客户端 110 以及服务端 120 的实现方式不做具体限定。

在一些实施例中, 参阅图 3B, 模型 200 可以为由多个网络层组成生成式模型。其中, 该多个网络层在生成式模型中的数据传递方向上依次排列, 也就是说, 一个网络层的输出是该网络层的一个网络层的输入。模块 210 包括该多个网络层中的至少一个网络层, 模块 220 包括该多个网络层中除该至少一个网络层之外的网络层。其中, 形成第一模块的网络层的输出是形成第二模块的网络层的输入。也就是说, 在生成式模型中的数据传递方向上, 模块 220 是模块 210 的下一层或多层。也就是说, 在该实施例中, 将组成生成式模型的多个网络层的一部分网络层部署到用户本地, 另一部分部署到云端。

其中, 生成式模型中的数据传递方向是指从数据输入至生成式模型到生成式模型输出数据这一过程中, 数据在生成式模型中的传递方向或者说前进方向。

在本申请实施例中, 生成式模型可以为文本到文本(text to text)的内容生成的模型、文本到图像(text to image)的内容生成的模型、文本到视频(text to video)的内容生成的模型或文本到音频的内容生成的模型。

模块 210 以内容生成指令为输入, 并输出编码向量。编码向量作为 AI 模型的中间处理结果, 难以用于推测 AI 模型的输入信息。也就是说, 编码向量为生成式模型基于内容生成指令生成目标内容过程中的中间结果, 难以用于推测输入至生成式模型的内容生成指令。具体而言, 生成式模型是基于深度学习训练得到的 AI 模型, 具有不可解释性。截至目前, 还没有能够利用生成式模型的中间结果或输出, 倒推出生成式模型的输入的技术。因此, 在该实施例中, 在用户本地, 利用生成式模型中的一部分网络层, 将用户发出的内容生成指令转换为生成式模型的中间结果, 并将中间结果发送至云端, 从而避免了将原始的内容生成指令直接发送至云端, 进而避免了敏感信息泄露。并且, 在该实施例中, 使用用户本地的计算资源对内容生成指令进行处理, 得到中间结果, 该中间结果直接作为部署在云端的模块 220 的输入,

从而无需云端对内容生成指令进行处理，节省了云端的计算资源，也提高了用户本地的计算资源的利用率。

在一些实施例中，参阅图 3C，模型 200 可以由过滤器和生成式模型组成。其中，生成式模型用作模块 220。过滤器也称为隐私过滤器，可以由用户设置。过滤器用于拦截或者过滤内容生成指令中的敏感信息，得到中间结果。此处的脱敏结果也称为脱敏后的内容生成指令。在一个示例中，过滤器包括用户设置的黑名单。黑名单包括用户列举的敏感信息。若内容生成指令包括黑名单中的信息，则过滤器拦截内容生成指令中黑名单上的信息。在一个示例中，过滤器包括用户设置的白名单。白名单包括用户列举的非敏感信息。若内容生成指令包括白名单之外的信息，则过滤器拦截内容生成指令中白名单之外的信息。在该实施例中，无需在用户本地运行对算力要求较高的神经网络，降低了对用户本地的算力要求。

上文示例介绍了本申请实施例提供的一种系统架构。接下来，结合该系统架构，对本申请实施例提供的一种内容生成方法进行介绍。其中，该方法具体可由内容生成系统 100 执行。如图 4 所示，该方法包括如下步骤。

步骤 401，客户端 110 接收用户发出的内容生成指令，所述内容生成指令用于指示所述模型 200 按照所述用户的意图生成目标内容。

内容生成指令包括该用户的意图。以文本到图像的内容生成为例，该用户想要一幅有关带烟筒的房子的图像，可以发出包含“画一幅带烟筒的房子”的内容生成指令。其中，“画一幅带烟筒的房子”为用户的意图。

其中，如图 3A 所示，客户端 110 处于用户本地，用户可以在本地将内容生成指令输入至客户端 110，保障了原始的内容生成指令不出用户本地，避免了隐私泄露。其中，原始的内容生成指令是指用户发出的、未经模块 210 或模型 200 处理的内容生成指令。

在一些实施例中，客户端 110 可以提供用户界面，该用户界面可供用户输入内容生成指令。例如，该用户界面可以包括编辑框，用户可以向编辑框中输入内容生成指令。

步骤 402，客户端 110 将所述内容生成指令输入至所述模块 210，使得所述模块 210 输出所述模型生成所述目标内容的中间结果。

如图 3A 所示，客户端 110 和模块 210 同处于用户本地，因此，可以在内容生成指令不出用户本地的情况下，将内容生成指令输入至模块 210，即将内容生成指令输入至模型 200，从而避免了用户隐私信息的泄露。

模块 210 可对内容生成指令进行数据脱敏处理，得到并输出中间结果。经过脱敏处理后的中间结果去除或保护了内容生成指令中的敏感信息。

在一些实施例中，如图 3B 所示，模型 200 为生成式模型，模块 210 为生成式模型中的至少一个网络层。该至少一个网络层可以对内容生成指令进行处理，得到生成式模型基于内容生成指令生成目标内容的中间结果。该多个网络层中除该至少一个网络层之外的网络层部署在云端，用作模块 220。

在一个示例中，该至少一个网络层可以包括输入层，该输出层可以对内容生成指令进行嵌入（embedding），得到内容生成指令的向量表示形式，即编码向量。该编码向量可以作为模块 210 输出的中间结果。

在一个示例中，该至少一个网络层还可以包括输入层和特征提取层。其中，特征提取层可以对内容生成指令的编码向量进行进一步计算，以提取的内容生成指令的特征，例如提取内容生成指令中用户的意图的特征，得到内容生成指令的特征向量。其中，该特征向量属于编码向量，为该至少一个网络层输出的编码向量。该特征向量可以作为模块 210 输出的中间结果。

在一个示例中，参阅图 5，作为模型 200 的生成式模型包括解释器和生成器。其中，解释器用于理解内容生成指令包含的用户意图，得到用户意图的编码向量。生成器用于按照用户意图的编码向量，利用生成器自身的知识以及掌握的信息，生成符合用户意图的内容。

在该示例中，解释器部署在用户本地，作为模块 210。生成器部署在云端，作为模块 220。内容生成指令包括用户的意图，例如“画一座带烟筒的房子”。作为模块 210 的解释器可以提取用户的意图的特征，得到用户的意图的特征向量。其中，提取用户的意图的特征也称为理解用户的意图，用户的意图的特征向量也可称为解释结果向量。模块 210 可以输出该解释结果向量。该解释结果向量作为步骤 402 得到的中间结果。其中，解释结果向量属于编码向量，具体为解释器输出的编码向量。

在一个示例中，参阅图 6，作为模型 200 的生成式模型包括解释器和生成器。其中，解释器可由多个网络层组成，且该多个网络层在生成式模型的数据传递方向上依次排列。解释器中的第一网络层可部署到用户本地，作为模块 210。解释器中的第二网络层和生成器可部署到云端，作为模块 220。其中，第一网络层为组成解释器的多个网络层中的至少一个网络层，第二网络层为该多个网络层中除第一网络层之外的网络层，并且第一网络层的输出为第二网络层的输入。其中，当第一网络层为两个及以上的网络层时，第一网络层的输出是指该两个及以上的网络层中在数据传递方向上的最后一个网络层的输出。当第二网络层为两个及以上的网络层时，第二网络层的输入是指该两个及以上的网络层中在数据传递方向上的第一个网络层的输入。

作为模块 210 的第一网络层可以提取用户的意图的特征，得到用户的意图的中间特征向量。用户的意图的中间特征向量也可称为中间解释结果向量。模块 210 可以输出该中间解释结果向量。该中间解释结果向量作为步骤 402 得到的中间结果。其中，中间解释结果向量属于编码向量，具体为第一网络层输出的编码向量。

在该示例中，部署在用户本地的模块 210 为部分解释器，在用户本地可完成内容生成指令的部分解释过程，剩余的解释过程以及全部的内容生成过程在云端进行，这样可以降低对用户本地的算力要求。举例而言，可以设定解释器为一个 100 层的神经网络，模块 210 可以包括该神经网络的前 50 层，该神经网络的后 50 层以及生成器部署在云端，如此，在用户本地，只需对内容生成指令进行前 50 层的数据处理，对用户本地的算力要求低。

在一个示例中，参阅图 7，作为模型 200 的生成式模型包括解释器和生成器。其中，生成器可由多个网络层组成，且该多个网络层在生成式模型的数据传递方向上依次排列。生成器中的第三网络层和解释器可部署到用户本地，作为模块 210。生成器中的第四网络层可部署到云端，作为模块 220。其中，第三网络层为组成生成器的多个网络层中的至少一个网络层，第四网络层为该多个网络层中除第三网络层之外的网络层，并且，第三网络层的输出为第四网络层的输入。其中，当第三网络层为两个及以上的网络层时，第三网络层的输出是指该两个及以上的网络层中在数据传递方向上的最后一个网络层的输出。当第四网络层为两个及以上的网络层时，第四网络层的输入是指该两个及以上的网络层中在数据传递方向上的第一个网络层的输入。

在该示例中，部署在用户本地的模块 210 包含解释器和部分生成器（即第三网络层）。其中，解释器可以提取用户的意图的特征，得到用户的意图的特征向量。该部分生成器可基于该特征向量，完成部分生成过程，得到部分生成结果。该部分生成结果作为步骤 402 得到的中间结果。该部分生成结果的表示形式也向量，即该部分生成结果也属于编码向量，为第三网络层输出的编码向量。

在该示例中，从用户本地发向云端的中间结果为部分生成结果，部分生成结果是对内容生成指令进行更多次运算而得到的，基于部分生成结果倒推内容生成指令的难度更大，从而有效避免了敏感信息泄露。

在一些实施例中，如图 3C 所示，模块 210 可以为部署在用户本地的过滤器，模块 220 为部署在云端的生成式模型。其中，模块 210 也可称为敏感信息过滤器，用于拦截或过滤内容生成指令中的敏感信息，得到脱敏后的内容生成指令。脱敏后的内容生成指令用作步骤 402 得到的中间结果。也就是说，模块 210 可以识别内容生成指令中的敏感信息，并拦截或过滤该敏感信息，使得该敏感信息不出用户本地。举例而言，如图 8 所示，可以设定用户发出的内容生成指令为“我的房子带烟筒，画一座带烟筒的房子”。其中，“我的房子带烟筒”包括用户的个人信息，属于敏感信息。模块 210 对该内容生成指令进行处理，得到的脱敏后的内容生成指令为“画一座带烟筒的房子”，不包含用户的敏感信息。

在一个示例中，模块 210 可以为由用户设置的白名单或黑名单，具体可以参考上文介绍，在此不再赘述。

回到图 4，在步骤 403，服务端 120 接收所述客户端 110 发送的所述中间结果，并将所述中间结果输入至所述模块 220，使得所述第二模块输出所述目标内容。

客户端 110 可以将模块 210 输出的中间结果，通过网络发送至位于云端的服务端 120。例如客户端 110 可以通过超文本传输安全协议（hypertext transfer protocol secure, HTTPS）或其他网络协议，将中间结果发送至服务端 120。

服务端 120 在接收到客户端 110 发送的中间结果时，可以将中间结果输入至模块 220。模块 220 可以基于该中间结果，得到目标内容。

当模块 210 为解释器，模块 220 为生成器时，中间结果为解释器输出的解释结果向量，生成器可以基于解释结果向量，生成目标内容。

当模块 210 为解释器中的第一网络层，模块 220 为解释器的第二网络层和生成器时，中间结果为中间解释结果向量，第二网络层可以基于中间解释结果，继续理解内容生成指令，得到解释结果向量，生成器可以基于解释结果向量，生成目标内容。

当模块 210 为解释器和生成器中的第三网络层，模块 220 为生成器中的第四网络层时，中间结果为部分生成结果，模块 220 可以基于该部分生成结果，继续生成内容，得到目标内容。

当模块 210 为过滤器，模块 220 为生成式模型时，中间结果为脱敏后的内容生成指令，生成式模型可基于脱敏后的内容生成指令，生成目标内容。

模块 220 在生成目标内容后，可输出目标内容。其中，目标内容可以为图片、视频、文字、音频中的任一种或多种。

在步骤 404，服务端 120 可以将模块 220 输出的目标内容，发送至客户端 110。其中，服务端 120 可以通过 HTTPS 或其他网络协议，将目标内容发送至客户端 110。

客户端 110 可以向用户提供目标内容，例如展示或播放目标内容，使得用户获得该目标内容。

在一些实施例中，模型 200 可以为生成式模型。生成式模型属于 AI 模型，需要通过训练才能获得。AI 模型的训练需要较高的算力和专业人员的操作。可以在云端训练模型 200。也就是说，模型 200 的训练可以由云厂商执行。其中，可以通过服务端 120 训练模型 200。在模型 200 训练完成后，即在模型 200 的损失函数（loss function）的值比较小（例如小于设定的值）后，服务端 220 可以获取模型 200 中的模块 210，并将模块 210 发送至用户本地，从而将模块 210 部署在用户本地。

在一个示例中，服务端 210 可以提供展示界面，展示界面用于展示模块 210 可下载。用户可以通过客户端 110 向服务端 210 发送下载请求。服务端 210 可以响应该下载请求，将模块 210 发送至用户本地。

在一个示例中，服务端 210 可以提供配置界面，配置界面用于配置模块 210。用户可通过客户端向服务端发送配置操作，该配置操作用于配置模块 210。服务端 210 可以基于该配置操作，配置模块 210，并将所配置的模块 210 发送至用户本地。例如配置操作为将解释器作为模块 210，服务端 210 基于该配置操作，将解释器作为模块 210 发送至用户本地。再例如，配置操作为将解释器中的第一网络层为模块 210，服务端 210 基于该配置操作，将第一网络层作为模块 210 发送至用户本地。再例如，配置操作为将解释器和生成器中的第三网络层为模块 210，服务端 210 基于该配置操作，将解释器和第三网络层作为模块 210 发送至用户本地。

在一些实施例中，用户可以对模块 210 的参数进行调整。如上所述，模块 210 可以包括至少一个网络层，其中，每个网络层包括多个神经元，每个神经元利用参数对数据进行计算，得到该层的处理结果。如此，用户对参数进行调整，使得云端或其他方无法感知模块 210 的参数，进而使得云端或其他方更加难以根据客户端发送的中间结果，倒推内容生成指令。

其中，用户可以发出参数调整操作，客户端 110 可以接收该参数调整操作，并基于参数调整操作，调整模块 210 的参数。

在一个例子中，用户可以通过客户端 110，利用记事本等应用，打开模块 210 的文件，直接修改模块 210 的参数。也就是说，用户发出的参数调整操作包括打开模块 210 的文件的操作和参数修改操作。客户端 110 可以响应打开模块 210 的文件的操作，打开模块 210 的文件，并响应参数修改操作，修改模块 210 的参数。

在一个例子中，可以在用户本地对模型 200 进行再训练，通过再训练，改变模块 210 的参数。具体而言，服务端 210 可以将模型 200 发送至用户本地。在用户本地，可以对模型 200 继续进行迭代训练，例如，迭代一次或多次，以修改模块 210 的参数。也就是说，用户发出的参数调整操作包括触发再训练的操作和终止再训练的操作。客户端 110 可以响应该触发再训练的操作，对模型 200 进行再训练，以及响应该终止再训练的操作，终止对模型的 200 的训练。

综合以上，在本申请实施例提供的内容生成方法，用于生成目标内容的模型一部分部署在用户本地，另一部分部署在云端。部署在用户本地的部分可以对内容生成指令进行数据脱敏等处理，得到难以用于倒推出内容生成指令的中间结果。然后，将中间结果发送至云端，使得云端基于中间结果生成目标内容，从而在无需用户本地运行全部模型的同时，避免了内容生成指令中的敏感信息泄露。

基于上文描述的内容，本申请实施例提供了一种内容生成系统 900。系统 900 用于利用模型生成内容。如图 9 所示，系统 900 包括客户端 910 和服务端 920，所述模型包括第一模块和第二模块，所述客户端 910 和所述第一模块部署在用户本地，所述服务端 920 和所述第二模块部署在云端，并且，所述第一模块的输出为所述第二模块的输入。其中，

所述客户端 910 用于接收用户发出的内容生成指令，所述内容生成指令用于指示所述模型按照所述用户的意图生成目标内容；

所述客户端 910 用于将所述内容生成指令输入至所述第一模块，使得所述第一模块输出所述模型生成所述目标内容的中间结果；

所述服务端 920 用于接收所述客户端 910 发送的所述中间结果，并将所述中间结果输入至所述第二模块，使得所述第二模块输出所述目标内容；

所述服务端 920 用于将所述目标内容发送至所述客户端 920。

在一些实施例中，所述模型为包括多个网络层的生成式模型，所述多个网络层在所述生成式模型的数据传递方向上依次排列，其中，所述第一模块包括所述多个网络层中的至少一个网络层，所述第二模块包括所述多个网络层中除所述至少一个网络层之外的网络层。

在一些实施例中，所述内容生成指令包括所述用户的意图，所述模型包括解释器和生成器，所述解释器用于提取所述用户的意图的特征，得到编码向量，所述生成器用于基于所述编码向量，生成所述目标内容；其中，所述第一模块包括所述解释器，所述第二模块包括所述生成器；或者，所述第一模块包括所述解释器中的第一网络层，所述第二模块包括所述解释器中的第二网络层和所述生成器；所述第一网络层的输出为所述第二网络层的输入；或者，所述第一模块包括所述解释器和所述生成器中的第三网络层，所述第二模块包括所述生成器中的第四网络层；所述第三网络层的输出为所述第四网络层的输入。

在一些实施例中，所述模型为生成式模型，所述服务端 920 还用于：训练所述模型；在所述模型训练完成后，将所述第一模块发送至所述用户本地。

在一些实施例中，所述客户端 910 还用于：接收用户发出的参数调整操作，并基于所述参数调整操作，调整所述第一模块的参数。

在一些实施例中，所述第一模块包括所述用户设置的敏感信息过滤器，所述过滤器用于拦截内容生成指令中的敏感信息，得到所述中间结果。

客户端 910、服务端 920 均可以通过软件实现，或者可以通过硬件实现。示例性的，接下来介绍客户端 910 的实现方式。类似的，服务端 920 的实现方式可以参考客户端 910 的实现方式。

装置作为软件功能单元的一种举例，客户端 910 可以包括运行在计算实例上的代码。其中，计算实例可以是物理主机（计算设备）、虚拟机、容器等计算设备中的至少一种。进一步地，上述计算设备可以是一台或者多台。例如，客户端 910 可以包括运行在多个主机/虚拟机/容器上的代码。需要说明的是，用于运行该应用程序的多个主机/虚拟机/容器可以分布在相同的区域（region）中，也可以分布在不同的 region 中。用于运行该代码的多个主机/虚拟机/容器可以分布在相同的可用区（availability zone, AZ）中，也可以分布在不同的 AZ 中，每个 AZ 包括一个数据中心或多个地理位置相近的数据中心。其中，通常一个 region 可以包括多个 AZ。

同样，用于运行该代码的多个主机/虚拟机/容器可以分布在同一个虚拟私有云（virtual private cloud, VPC）中，也可以分布在多个 VPC 中。其中，通常一个 VPC 设置在一个 region 内。同一 region 内两个 VPC 之间，以及不同 region 的 VPC 之间跨区通信需在每个 VPC 内设置通信网关，经通信网关实现 VPC 之间的互连。

装置作为硬件功能单元的一种举例，客户端 910 可以包括至少一个计算设备，如服务器、终端等。或者，客户端 910 也可以是利用专用集成电路（application-specific integrated circuit, ASIC）实现、或可编程逻辑器件（programmable logic device, PLD）实现的设备等。其中，上述 PLD 可以是复杂程序逻辑器件（complex programmable logical device, CPLD）、现场可编程门阵列（field-programmable gate array, FPGA）、通用阵列逻辑（generic array logic, GAL）或其任意组合实现。

客户端 910 包括的多个计算设备可以分布在相同的 region 中，也可以分布在不同的 region 中。客户端 910 包括的多个计算设备可以分布在相同的 AZ 中，也可以分布在不同的 AZ 中。同样，客户端 910 包括的多个计算设备可以分布在同一个 VPC 中，也可以分布在多个 VPC 中。其中，所述多个计算设备可以是服务器、ASIC、PLD、CPLD、FPGA 和 GAL 等计算设备的任意组合。

基于上文描述的内容，本申请实施例提供了一种利用模型生成内容的方法。该方法可应用于部署在用户本地的客户端，所述客户端所在的内容生成系统还包括部署在云端的服务端，所述模型包括第一模块和第二模块，所述第一模块部署在所述用户本地，所述第二模块部署在所述云端，并且，所述第一模块的输出为所述第二模块的输入。如图 10 所示，该方法包括如下步骤：

步骤 1001，所述客户端接收用户发出的内容生成指令，所述内容生成指令用于指示所述模型按照所述用户的意图生成目标内容；

步骤 1002，所述客户端将所述内容生成指令输入至所述第一模块，使得所述第一模块输出所述模型生成所述目标内容的中间结果；

步骤 1003，所述客户端将所述中间结果发送至所述服务端；其中，所述服务端用于将所述中间结果输入至所述第二模块，使得所述第二模块输出所述目标内容；

步骤 1004，所述客户端接收所述服务端发送的所述目标内容。

在一些实施例中，所述方法还包括：所述客户端接收用户发出的参数调整操作，并基于所述参数调整操作，调整所述第一模块的参数。

参阅图 11，本申请实施例还提供了一种利用模型生成内容的客户端 1100。客户端 1100 部署在用户本地，所述客户端 1100 所在的内容生成系统还包括部署在云端的服务端，所述模型包括第一模块和第二模块，所述第一模块部署在所述用户本地，所述第二模块部署在所述云端，并且，所述第一模块的输出为所述第二模块的输入。所述客户端 1100 包括：

接收模块 1110，用于接收用户发出的内容生成指令，所述内容生成指令用于指示所述模型按照所述用户的意图生成目标内容；

输入模块 1120，用于将所述内容生成指令输入至所述第一模块，使得所述第一模块输出所述模型生成所述目标内容的中间结果；

通信模块 1130，用于将所述中间结果发送至所述服务端；其中，所述服务端用于将所述中间结果输入至所述第二模块，使得所述第二模块输出所述目标内容；

所述通信模块 1130 还用于接收所述服务端发送的所述目标内容。

其中，接收模块 1110、输入模块 1120 和通信模块 1130 均可以通过软件实现，或者可以通过硬件实现。示例性的，接下来以接收模块 1110 为例，介绍接收模块 1110 的实现方式。类似的，输入模块 1120 和通信模块 1130 的实现方式可以参考接收模块 1110 的实现方式。

模块作为软件功能单元的一种举例，接收模块 1110 可以包括运行在计算实例上的代码。其中，计算实例可以包括物理主机（计算设备）、虚拟机、容器中的至少一种。进一步地，上述计算实例可以是一台或者多台。例如，接收模块 1110 可以包括运行在多个主机/虚拟机/容器上的代码。需要说明的是，用于运行该代码的多个主机/虚拟机/容器可以分布在相同的 region 中，也可以分布在不同的 region 中。进一步地，用于运行该代码的多个主机/虚拟机/容器可以分布在相同的 AZ 中，也可以分布在不同的 AZ 中，每个 AZ 包括一个数据中心或多个地理位置相近的数据中心。其中，通常一个 region 可以包括多个 AZ。

同样，用于运行该代码的多个主机/虚拟机/容器可以分布在同一个 VPC 中，也可以分布在多个 VPC 中。其中，通常一个 VPC 设置在一个 region 内，同一 region 内两个 VPC 之间，以及不同 region 的 VPC 之间跨区通信需在每个 VPC 内设置通信网关，经通信网关实现 VPC 之间的互连。

模块作为硬件功能单元的一种举例，接收模块 1110 可以包括至少一个计算设备，如服务器等。或者，接收模块 1110 也可以是利用 ASIC 实现、或可编程逻辑器件 PLD 实现的设备等。其中，上述 PLD 可以是 CPLD、FPGA、GAL 或其任意组合实现。

接收模块 1110 包括的多个计算设备可以分布在相同的 region 中，也可以分布在不同的 region 中。接收模块 1110 包括的多个计算设备可以分布在相同的 AZ 中，也可以分布在不同的 AZ 中。同样，接收模块 1110 包括的多个计算设备可以分布在同一个 VPC 中，也可以分布在多个 VPC 中。其中，所述多个计算设备可以是服务器、ASIC、PLD、CPLD、FPGA 和 GAL 等计算设备的任意组合。

需要说明的是，在其他实施例中，接收模块 1110 可以用于执行图 10 所示方法中的任意步骤，输入模块 1120 可以用于执行图 10 所示方法中的任意步骤，通信模块 1130 可以用于执行图 10 所示方法中的任意步骤。接收模块 1110、输入模块 1120 和通信模块 1130 负责实现的步骤可根据需要指定，通过接收

模块 1110、输入模块 1120 和通信模块 1130 分别实现图 10 所示方法中不同的步骤来实现客户端 1100 的全部功能。

本申请还提供一种计算设备 1200。如图 12 所示，计算设备 1200 包括：总线 1202、处理器 1204、存储器 1206 和通信接口 1208。处理器 1204、存储器 1206 和通信接口 1208 之间通过总线 1202 通信。计算设备 1200 可以是服务器或终端设备。应理解，本申请不限定计算设备 1200 中的处理器、存储器的个数。

总线 1202 可以是外设部件互连标准 (peripheral component interconnect, PCI) 总线或扩展工业标准结构 (extended industry standard architecture, EISA) 总线等。总线可以分为地址总线、数据总线、控制总线等。为便于表示，图 12 中仅用一条线表示，但并不表示仅有一根总线或一种类型的总线。总线 1202 可包括在计算设备 1200 各个部件 (例如，存储器 1206、处理器 1204、通信接口 1208) 之间传递信息的通路。

处理器 1204 可以包括中央处理器 (central processing unit, CPU)、图形处理器 (graphics processing unit, GPU)、微处理器 (micro processor, MP) 或者数字信号处理器 (digital signal processor, DSP) 等处理器中的任意一种或多种。

存储器 1206 可以包括易失性存储器 (volatile memory)，例如随机存取存储器 (random access memory, RAM)。存储器 1206 还可以包括非易失性存储器 (non-volatile memory)，例如只读存储器 (read-only memory, ROM)，快闪存储器，机械硬盘 (hard disk drive, HDD) 或固态硬盘 (solid state drive, SSD)。

存储器 1206 中存储有可执行的程序代码，处理器 1204 执行该可执行的程序代码以分别实现前述接收模块 1110、输入模块 1120 和通信模块 1130 的功能，从而实现图 10 所示方法。也即，存储器 1206 上存有用以执行图 10 所示方法的指令。

通信接口 1208 使用例如但不限于网络接口卡、收发器一类的收发模块，来实现计算设备 1200 与其他设备或通信网络之间的通信。

本申请实施例还提供了一种计算设备集群。该计算设备集群包括至少一台计算设备。该计算设备可以是服务器，例如是边缘服务器，或者是本地数据中心中的本地服务器。在一些实施例中，计算设备也可以是台式机、笔记本电脑或者智能手机等终端设备。

如图 13 所示，所述计算设备集群包括至少一个计算设备 1200。计算设备集群中的一个或多个计算设备 1200 中的存储器 1206 中可以存有相同的用于执行图 10 所示方法的指令。

在一些可能的实现方式中，该计算设备集群中的一个或多个计算设备 1200 的存储器 1206 中也可以分别存有用以执行图 10 所示方法的部分指令。换言之，一个或多个计算设备 1200 的组合可以共同执行用于执行图 10 所示方法的指令。

需要说明的是，计算设备集群中的不同的计算设备 1200 中的存储器 1206 可以存储不同的指令，分别用于执行客户端 1100 的部分功能。也即，不同的计算设备 1200 中的存储器 1206 存储的指令可以实现接收模块 1110、输入模块 1120 和通信模块 1130 中的一个或多个模块的功能。

在一些可能的实现方式中，计算设备集群中的一个或多个计算设备可以通过网络连接。其中，所述网络可以是广域网或局域网等等。图 14 示出了一种可能的实现方式。如图 14 所示，两个计算设备 1200A 和 1200B 之间通过网络进行连接。具体地，通过各个计算设备中的通信接口与所述网络进行连接。在这一类可能的实现方式中，计算设备 1200A 中的存储器 1206 中存有执行接收模块 1110 的功能的指令。同时，计算设备 1200B 中的存储器 1206 中存有执行输入模块 1120 和通信模块 1130 的功能的指令。

应理解，图 14 中示出的计算设备 1200A 的功能也可以由多个计算设备 1200 完成。同样，计算设备 1200B 的功能也可以由多个计算设备 1200 完成。

本申请实施例还提供了另一种计算设备集群。该计算设备集群中各计算设备之间的连接关系可以类似的参考图 13 和图 14 所述计算设备集群的连接方式。不同的是，该计算设备集群中的一个或多个计算设备 1200 中的存储器 1206 中可以存有相同的用于执行图 10 所示方法的指令。

在一些可能的实现方式中，该计算设备集群中的一个或多个计算设备 1200 的存储器 1206 中也可以分别存有用以执行图 10 所示方法的部分指令。换言之，一个或多个计算设备 1200 的组合可以共同执行用于执行图 10 所示方法的指令。

本申请实施例还提供了一种包含指令的计算机程序产品。所述计算机程序产品可以是包含指令的，能够运行在计算设备上或被储存在任何可用介质中的软件或程序产品。当所述计算机程序产品在至少一个计算设备上运行时，使得至少一个计算设备执行图 10 所示方法。

本申请实施例还提供了一种计算机可读存储介质。所述计算机可读存储介质可以是计算设备能够存储的任何可用介质或者是包含一个或多个可用介质的数据中心等主机迁移设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘）等。该计算机可读存储介质包括指令，所述指令指示计算设备执行图 10 所示方法。

最后应说明的是：以上实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的保护范围。

权 利 要 求 书

1. 一种利用模型生成内容的方法，其特征在于，应用于内容生成系统，所述系统包括客户端和服务端，所述模型包括第一模块和第二模块，所述客户端和所述第一模块部署在用户本地，所述服务端和所述第二模块部署在云端，并且，所述第一模块的输出为所述第二模块的输入；所述方法包括：

所述客户端接收用户发出的内容生成指令，所述内容生成指令用于指示所述模型按照所述用户的意图生成目标内容；

所述客户端将所述内容生成指令输入至所述第一模块，使得所述第一模块输出所述模型生成所述目标内容的中间结果；

所述服务端接收所述客户端发送的所述中间结果，并将所述中间结果输入至所述第二模块，使得所述第二模块输出所述目标内容；

所述服务端将所述目标内容发送至所述客户端。

2. 根据权利要求 1 所述的方法，其特征在于，所述模型为包括多个网络层的生成式模型，所述多个网络层在所述生成式模型的数据传递方向上依次排列，其中，所述第一模块包括所述多个网络层中的至少一个网络层，所述第二模块包括所述多个网络层中除所述至少一个网络层之外的网络层。

3. 根据权利要求 1 或 2 所述的方法，其特征在于，所述内容生成指令包括所述用户的意图，所述模型包括解释器和生成器，所述解释器用于提取所述用户的意图的特征，得到编码向量，所述生成器用于基于所述编码向量，生成所述目标内容；其中，

所述第一模块包括所述解释器，所述第二模块包括所述生成器；

或者，

所述第一模块包括所述解释器中的第一网络层，所述第二模块包括所述解释器中的第二网络层和所述生成器；所述第一网络层的输出为所述第二网络层的输入；

或者，

所述第一模块包括所述解释器和所述生成器中的第三网络层，所述第二模块包括所述生成器中的第四网络层；所述第三网络层的输出为所述第四网络层的输入。

4. 根据权利要求 1-3 中任一项所述的方法，其特征在于，所述模型为生成式模型，所述方法还包括：所述服务端训练所述模型；

在所述模型训练完成后，所述服务端将所述第一模块发送至所述用户本地。

5. 根据权利要求 1-4 中任一项所述的方法，其特征在于，所述方法还包括：所述客户端接收用户发出的参数调整操作，并基于所述参数调整操作，调整所述第一模块的参数。

6. 根据权利要求 1 所述的方法，其特征在于，所述第一模块包括所述用户设置的敏感信息过滤器，所述过滤器用于拦截内容生成指令中的敏感信息，得到所述中间结果。

7. 一种内容生成系统，其特征在于，所述系统用于利用模型生成内容，其中，所述系统包括客户端和服务端，所述模型包括第一模块和第二模块，所述客户端和所述第一模块部署在用户本地，所述服务端和所述第二模块部署在云端，并且，所述第一模块的输出为所述第二模块的输入；

所述客户端用于接收用户发出的内容生成指令，所述内容生成指令用于指示所述模型按照所述用户的意图生成目标内容；

所述客户端用于将所述内容生成指令输入至所述第一模块，使得所述第一模块输出所述模型生成所述目标内容的中间结果；

所述服务端用于接收所述客户端发送的所述中间结果，并将所述中间结果输入至所述第二模块，使得所述第二模块输出所述目标内容；

所述服务端用于将所述目标内容发送至所述客户端。

8. 根据权利要求 7 所述的系统，其特征在于，所述模型为包括多个网络层的生成式模型，所述多个网络层在所述生成式模型的数据传递方向上依次排列，其中，所述第一模块包括所述多个网络层中的至少一个网络层，所述第二模块包括所述多个网络层中除所述至少一个网络层之外的网络层。

9. 根据权利要求 7 或 8 所述的系统，其特征在于，所述内容生成指令包括所述用户的意图，所述模型包括解释器和生成器，所述解释器用于提取所述用户的意图的特征，得到编码向量，所述生成器用于基于所述编码向量，生成所述目标内容；其中，

所述第一模块包括所述解释器，所述第二模块包括所述生成器；

或者，

所述第一模块包括所述解释器中的第一网络层，所述第二模块包括所述解释器中的第二网络层和所述生成器；所述第一网络层的输出为所述第二网络层的输入；

或者，

所述第一模块包括所述解释器和所述生成器中的第三网络层，所述第二模块包括所述生成器中的第四网络层；所述第三网络层的输出为所述第四网络层的输入。

10. 根据权利要求 7-9 中任一项所述的系统，其特征在于，所述模型为生成式模型，所述服务端还用于：

训练所述模型；

在所述模型训练完成后，将所述第一模块发送至所述用户本地。

11. 根据权利要求 7-10 中任一项所述的系统，其特征在于，所述客户端还用于：接收用户发出的参数调整操作，并基于所述参数调整操作，调整所述第一模块的参数。

12. 根据权利要求 7 所述的系统，其特征在于，所述第一模块包括所述用户设置的敏感信息过滤器，所述过滤器用于拦截内容生成指令中的敏感信息，得到所述中间结果。

13. 一种利用模型生成内容的方法，其特征在于，应用于部署在用户本地的客户端，所述客户端所在的内容生成系统还包括部署在云端的服务端，所述模型包括第一模块和第二模块，所述第一模块部署在所述用户本地，所述第二模块部署在所述云端，并且，所述第一模块的输出为所述第二模块的输入；所述方法包括：

所述客户端接收用户发出的内容生成指令，所述内容生成指令用于指示所述模型按照所述用户的意图生成目标内容；

所述客户端将所述内容生成指令输入至所述第一模块，使得所述第一模块输出所述模型生成所述目标内容的中间结果；

所述客户端将所述中间结果发送至所述服务端；其中，所述服务端用于将所述中间结果输入至所述第二模块，使得所述第二模块输出所述目标内容；

所述客户端接收所述服务端发送的所述目标内容。

14. 根据权利要求 13 所述的方法，其特征在于，所述方法还包括：所述客户端接收用户发出的参数调整操作，并基于所述参数调整操作，调整所述第一模块的参数。

15. 一种利用模型生成内容的客户端，其特征在于，所述客户端部署在用户本地，所述客户端所在的内容生成系统还包括部署在云端的服务端，所述模型包括第一模块和第二模块，所述第一模块部署在所述用户本地，所述第二模块部署在所述云端，并且，所述第一模块的输出为所述第二模块的输入；所述客户端包括：

接收模块，用于接收用户发出的内容生成指令，所述内容生成指令用于指示所述模型按照所述用户的意图生成目标内容；

输入模块，用于将所述内容生成指令输入至所述第一模块，使得所述第一模块输出所述模型生成所述目标内容的中间结果；

通信模块，用于将所述中间结果发送至所述服务端；其中，所述服务端用于将所述中间结果输入至所述第二模块，使得所述第二模块输出所述目标内容；

所述通信模块还用于接收所述服务端发送的所述目标内容。

16. 根据权利要求 15 所述的客户端，其特征在于，所述接收模块还用于：接收用户发出的参数调整操作，并基于所述参数调整操作，调整所述第一模块的参数。

17. 一种计算设备集群，其特征在于，包括至少一个计算设备，每个计算设备包括处理器和存储器；

所述至少一个计算设备的处理器用于执行所述至少一个计算设备的存储器中存储的指令，以使得所述计算设备集群执行如权利要求 13 或 14 所述的方法。

18. 一种包含指令的计算机程序产品，其特征在于，当所述指令被计算设备集群运行时，使得所述计算设备集群执行如权利要求 13 或 14 所述的方法。

19. 一种计算机可读存储介质，其特征在于，包括计算机程序指令，当所述计算机程序指令由计算设备集群执行时，所述计算设备集群执行如权利要求 13 或 14 所述的方法。