

## 《人工智能拟人化互动服务管理暂行办法（征求意见稿）》速评

作者：贾维恒 | 龙继伟

2025年12月27日，国家互联网信息办公室（下称“网信办”）发布《人工智能拟人化互动服务管理暂行办法（征求意见稿）》（下称“《征求意见稿》”），首次从监管的角度构筑对人工智能拟人化互动服务的监管框架。本文将简要介绍《征求意见稿》的几个要点，并从法律实务角度进行解读，供各位读者参考。

### 一、上位法和适用范围

《征求意见稿》第一条规定，其上位法包括《民法典》、《网络安全法》、《数据安全法》、《科学技术进步法》、《个人信息保护法》、《网络数据安全条例》、《未成年人网络保护条例》、《互联网信息服务管理办法》等法律、行政法规。相对于欧盟等国家和地区已经对人工智能出台了诸多基本的监管法律，我国人工智能产业发展迅猛，但尚无专门针对人工智能的法律层面专项立法，所以对于人工智能拟人化服务的监管，也不得不借助其他相关的基本法律。另外，据媒体报道，由于拟人化服务已经出现了一些极端的事件，也加快了政府对于拟人化人工智能服务的监管需求。

《征求意见稿》第二条规定，《征求意见稿》的适用范围为“利用人工智能技术，向中华人民共和国境内公众提供模拟人类人格特征、思维模式和沟通风格，通过文字、图片、音频、视频等方式与人类进行情感互动的产品或者服务（以下称拟人化互动服务）”。在适用范围上至少有如下几点值得关注：（1）适用主体是拟人化互动服务提供者（下称“提供者”），没有限定为中国境内或境外主体，亦未限定为法人或非法人；（2）适用的行为是向中国境内提供拟人化互动服务，在文义上可以理解为在中国境内提供，或者从中国境外向境内提供；（3）“拟人化互动服务”的落脚点是“情感互动”，在界定和解释上或许仍有进一步明确的空间（例如，纯粹的知识类、科普类互动，是否被排除在外）。

### 二、拟人化互动服务行为红线

《征求意见稿》第七条以不完全列举的形式对提供和使用拟人化互动服务过程中不得开展的活动（下称“行为红线”）进行规定。行为红线包括：（1）危害国家安全、社会秩序相关；（2）“黄、赌、毒”相关；（3）侮辱、诽谤相关；（4）虚假承诺相关；（5）自杀、自残、心理健康相关；（6）不合理决策相关；（7）涉敏感信息；（8）其他。

可以认为，行为红线的明确对于提供者进一步优化拟人化互动服务提供过程中生成的内容提出了更高要求。例如，在通过图片方式进行互动的场景下，如何认定图片的颜色、某些图片要素是否具有特殊含义，这将对进一步判断该图片是否存在明示或暗示的违规内容产生影响。

### 三、适用的行政手续

《征求意见稿》中涉及的主要行政手续包括：

#### （一）安全评估

《征求意见稿》第二十一条规定，提供者涉及如下情形之一时，应当向省级网信部门提交评估报告，开展安全评估：(i)上线或增设拟人化互动服务的功能；(ii)使用新技术新应用，导致拟人化互动服务发生重大变更；(iii)注册用户达 100 万以上或者月活跃用户达 10 万以上；(iv)可能影响国家安全、公共利益等；以及，(v)其他情形。我们理解，前述情形中，有的情形较为容易判断（例如，相关功能的上线或增设，以及用户数量是否达到门槛），而有的则具有进一步解释空间（例如，何谓服务发生重大变更，何谓可能影响国家安全）。《征求意见稿》未规定履行安全评估义务的豁免情形。

#### （二）重大安全隐患报告

《征求意见稿》第二十三条规定，提供者发现用户存在重大安全隐患的，应当采取限制功能、暂停或者终止提供服务，保存有关记录，并向有关主管部门报告。从实务角度看，如何解读“重大安全隐患”的内涵和外延，将在很大程度上决定提供者是否应当履行重大安全隐患报告义务。

#### （三）算法备案

《征求意见稿》第二十五条规定，提供者应当按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续。根据《互联网信息服务算法推荐管理规定》第二十四条，应当履行备案手续的主体为“具有舆论属性或者社会动员能力的算法推荐服务提供者”。我们理解，在《征求意见稿》和《互联网信息服务算法推荐管理规定》重合监管的部分，提供者应当分别履行相应的义务。

#### （四）未成年人信息处理情况的合规审计

《征求意见稿》第十五条第二款规定，提供者应当自行或者委托专业机构每年对其处理未成年人个人信息遵守法律、行政法规的情况进行合规审计。

### 四、特殊用户人群的服务提供规则

《征求意见稿》针对未成年人、老年人等特殊用户人群，分别增设了相应的服务提供规则。

#### （一）未成年人

对于面向未成年人用户的服务提供规则主要包括：(i)建立未成年人模式，明确在提供情感陪伴服务时应当取得监护人的明确同意；以及(ii)未成年人模式下收集数据向第三方提供时，还需取得监护人单独同意。值得注意的是，《个人信息保护法》项下对于未成年人个人信息的特殊处理规则中有“不满十四周岁”的限制，而《征求意见稿》则没有相关限制，这体现了对于未成年人群体的特别保护。

#### （二）老年人

对于面向老年人用户的服务提供规则主要包括：(i)设置服务紧急联系人，并在适用情形下紧急通知紧急联系人；以及(ii)明确不得提供模拟老年人用户亲属、特定关系人的服务。

## 五、数据和个人信息处理规则

《征求意见稿》对于数据、个人信息的处理提出如下规则：

### （一）训练数据的管理

《征求意见稿》第十条详细规定了在预训练、优化训练等数据处理活动场景下的训练数据管理规定，例如，应当对训练数据开展清洗、标注，评估合成数据安全性等。

### （二）用户交互数据的安全保护

《征求意见稿》第十四条规定，提供者应当：(i)采取数据加密、安全审计、访问控制等措施；(ii)除非取得相关主体同意或其他豁免情形下，不得向第三方提供用户交互数据；以及(iii)向用户提供删除交互数据的选项。

## 六、法律责任

在法律责任方面，根据《征求意见稿》第二十九条，除了依照现行法律、行政法规进行处罚之外，提供者面临的法律责任主要包括警告、通报批评、责令限期整改、责令暂停提供服务。我们注意到，《征求意见稿》中没有单独规定罚金、没收违法所得的罚则。后续的生效文本中是否有调整，我们将持续关注。

## 小结

《征求意见稿》以“拟人化互动服务”为切入点，对于提供者增加了安全评估和报备的强制义务，和违反相关义务而受到行政处罚的风险。其直接影响是提供者的成本可能会因监管而增加，内容的自由度也会受到限制和监管。也就是说，人工智能说了不该说的话，出了违规的主意，监管机构可以对提供者进行处罚。

## 特别声明

汉坤律师事务所编写《汉坤法律评述》的目的仅为帮助客户及时了解中国或其他相关司法管辖区法律及实务的最新动态和发展，仅供参考，不应被视为任何意义上的法律意见或法律依据。

如您对本期《汉坤法律评述》内容有任何问题或建议，请与汉坤律师事务所以下人员联系：

### 贾维恒

电话： +86 21 6080 0396

Email: [weiheng.jia@hankunlaw.com](mailto:weiheng.jia@hankunlaw.com)