

团 体 标 准

T/CESA 1041—2019

全国团体标准信息平台

信息技术 人工智能 服务能力成熟度评价 参考模型

Information technology- Artificial intelligence- Reference model of service capability
maturity evaluation

全国团体标准信息平台

2019 - 04 - 01 发布

2019 - 04 - 01 实施

中国电子工业标准化技术协会

发 布

目 次

前 言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 参考模型	2
6 成熟度等级	2
6.1 成熟度等级划分	2
6.2 成熟度分级规则	3
7 能力框架	4
7.1 能力框架结构	4
7.2 能力项描述	5
附录 A （规范性附录） 成熟度参考模型评分	8
附录 B （规范性附录） 能力子项权重及成熟度参考模型评分	9

前 言

本部分按照GB/T 1.1—2009《标准化工作导则 第1部分：标准的结构和编写》给出的规则起草。请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由中国电子技术标准化研究院提出并归口。

本标准起草单位：中国电子技术标准化研究院、小米通讯技术有限公司、苏州思必驰信息科技有限公司、中国航空综合技术研究所、华为技术有限公司、山东省计算中心（国家超级计算济南中心）、中国医学科学院生物医学工程研究所、浪潮软件集团有限公司、宁波GQY视讯股份有限公司、曙光信息产业（北京）有限公司、上海智能制造系统创新中心有限公司、电子科技大学、深圳市人马互动科技有限公司、工业和信息化部电子第五研究所、广州广电运通金融电子股份有限公司、玉养信息科技有限公司（上海）有限公司、华夏芯（北京）通用处理器技术有限公司、南京行者易智能交通科技有限公司、深圳市长亮科技股份有限公司、威麟信息技术开发（上海）有限公司、中科博宏（北京）科技有限公司、贵州电网有限责任公司信息中心等。

本标准主要起草人：代红、董建、张群、马珊珊、周珏嘉、朱亚军、赵群、汪小娟、王燕妮、郭涛涛、王洁萍、符海芳、光亮、高永超、李刚、徐圣普、蒲江波、黄先芝、王功明、陈养彬、王建宽、孙国忠、吕灼恒、王飞、谭李诺、罗光春、张栗粽、胡上峰、靳志业、黄林轶、徐鹏飞、林冠辰、陈良旭、朱兆颖、刘军、李军、辜敏、曾曦、翁家良、翁斌、王皓然等。

全国团体标准信息平台

信息技术 人工智能 服务能力成熟度评价参考模型

1 范围

本标准提出了人工智能服务能力成熟度评价参考模型，规定了成熟度等级、能力框架和评价方法。本标准适用于对云服务提供商提供的人工智能服务能力的成熟度评估，以及服务能力成熟度模型中某项能力主域、能力子域的单项评估。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 5271.28-2001	信息技术	词汇	第28部分:人工智能	基本概念与专家系统
GB/T 5271.29-2006	信息技术	词汇	第29部分:人工智能	语音识别与合成
GB/T 5271.31-2006	信息技术	词汇	第31部分:人工智能	机器学习
GB/T 5271.34-2006	信息技术	词汇	第34部分:人工智能	神经网络

3 术语和定义

GB/T 5271.28-2001、GB/T 5271.29-2006、GB/T 5271.31-2006、GB/T 5271.34-2006界定的以及下列术语和定义适用于本文件。

3.1

能力子项 *capability sub-item*

能力子域中构成成熟度的若干特性指标单项。

3.2

能力子项权重 *factor of capability sub-item*

对应能力子项在所述能力子域评价中的权重因子。

3.3

服务可用性 *service availability*

服务客户发起服务请求后，服务可访问的时间占总服务时间的比例。

注：可用性的计算是在一系列预定义的时间段中，服务可用时间之和占预定义时间段之和的比例，可排除允许的服务不可用时间。

4 缩略语

AI：人工智能（Artificial Intelligence）

5 参考模型

服务能力成熟度评价参考模型见图1，包括能力框架、评价指标体系、评价算法和成熟度等级。

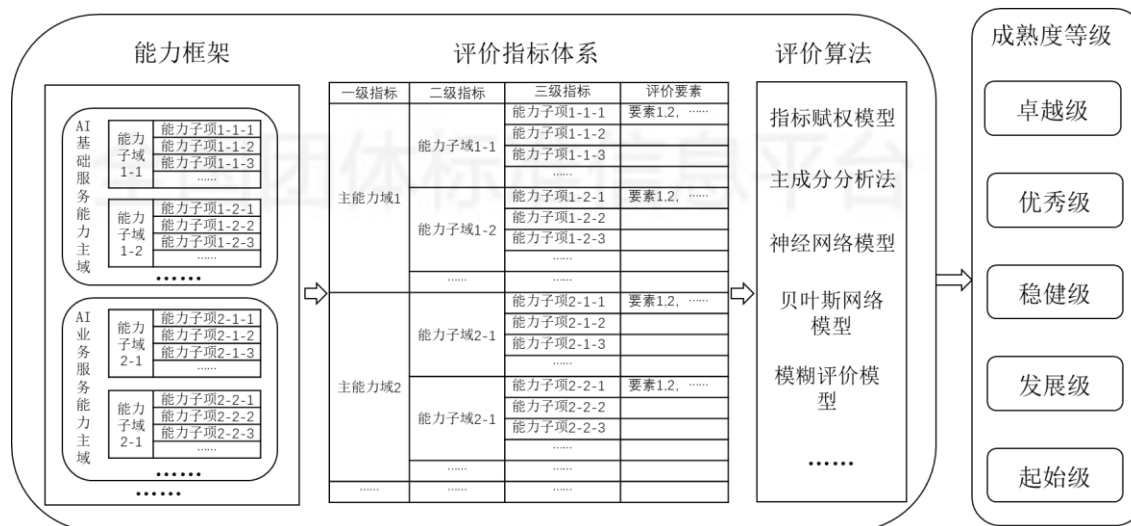


图1 人工智能服务能力成熟度评价参考模型

能力框架由多个能力主域组成，每个能力主域代表成熟度等级评价中主要的领域划分，由一组相关的能力子域组成。每个能力子域代表该能力主域中构成成熟度水平的若干技术能力领域，由一组相关的能力子项组成。

评价指标体系由三级指标构成，分别与能力主域、能力子域和能力子项对应，每个能力子项代表该能力子域构成成熟度的若干特性指标。

评价算法是指在获得各项特性指标数据后，选择某一方法实现对评价指标体系的综合计算，其输出结果为某一成熟度等级。评价算法包括且不限于指标赋权模型、主成分分析法、神经网络模型、贝叶斯网络模型、模糊评价模型。指标赋权模型的评价参见附录B。

成熟度等级是从基础服务能力、业务服务能力等方面对人工智能服务能力的分级评价结果。

6 成熟度等级

6.1 成熟度等级划分

人工智能服务能力成熟度等级分为5级：起始级、发展级、稳健级、优秀级、卓越级，并用一、二、三、四、五表示，如图2。

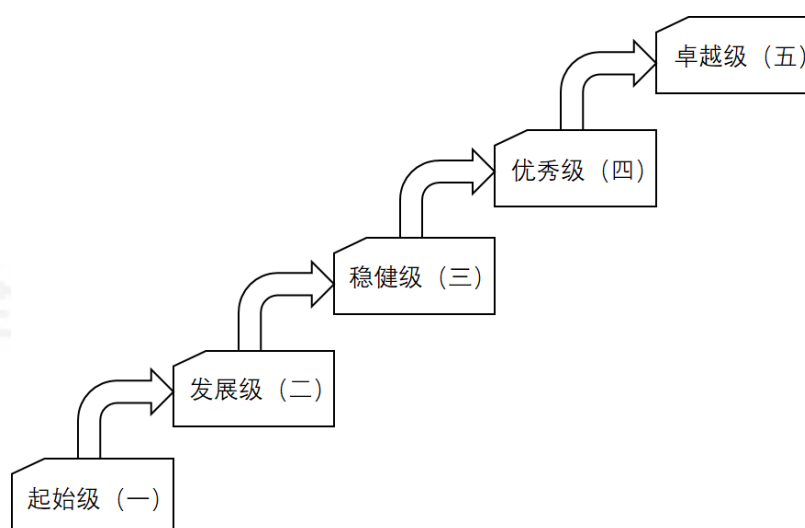


图2 人工智能服务能力成熟度等级

每个成熟度等级表明人工智能服务能力所达到的成熟度水平。

人工智能服务能力的改进和成熟度提升是通过渐进的方式来实现的,较高的成熟度等级涵盖了低于其成熟度等级的全部要求。

6.2 成熟度分级规则

人工智能服务能力、某项能力主域、能力子域成熟度分级规则见表1。

表1 人工智能服务能力成熟度分级规则

等 级	特 征
起始级	能力主域中各个能力子域的能力子项只达到较低或最低的要求。 人工智能服务只能完成基本的功能,在技术水平、系统性能、可扩展性、安全性等方面只能达到较低或最低的要求。
发展级	能力主域中大多数能力子域的能力子项只达到较低的要求。 人工智能服务能够完成基本的功能,在技术水平、系统性能、可扩展性、安全性等方面中的一部分可以达到较低或中等的要求。
稳健级	能力主域中大多数能力子域的能力子项达到中等的要求。 人工智能服务能够完善的完成基本的功能,在技术水平、系统性能、可扩展性、安全性等方面中的大部分可以达到或超过中等的要求。
优秀级	能力主域中大多数能力子域的能力子项能达到中等或较高的要求。 人工智能服务除了有优秀的功能实现,在技术水平、系统性能、可扩展性、安全性等方面中的大部分可以达到较高的要求。
卓越级	能力主域中大多数能力子域的能力子项能达到较高或最高的要求。 人工智能服务除了有卓越的功能实现,在技术水平、系统性能、可扩展性、安全性等方面中的大部分可以达到最高的要求。

服务能力成熟度分级可采用评分量化的方式进行,其计算方法参见附录A。

7 能力框架

7.1 能力框架结构

人工智能服务能力成熟度框架应分为AI基础服务能力、AI业务服务能力。各能力主域的能力子域和能力子项按照衡量相关能力所需的技术水平、系统性能、可扩展性、安全性等维度进行规定，其框架见表2。

表2 人工智能服务能力成熟度能力框架

能力主域	能力子域	能力子项
AI基础服务能力	存储	数据管理
		多样性
		分布式可扩展
		其他
	计算	数据计算
		集群效率
		稳定性
		可扩展性
		灵活性
		统计监控
	安全与隐私	协议安全
		数据隐私和算法安全
		数据加密
		网络隔离
		权限管理和鉴权
		行为审计
		恢复备份
	机器学习	全生命周期
		训练性能
兼容性		
灵活性		
易用性		
算法库		
AI业务服务能力	智能语音	语音识别
		语音合成
		声纹识别
	自然语言理解	词法、句法分析
		信息抽取
		信息检索
		语义推理
		情感分析

表2 人工智能服务能力成熟度能力框架（续）

能力主域	能力子域	能力子项	
	智能对话	知识图谱	
		闲聊	
		知识型问答	
		任务导向对话	
		情感对话	
		启发式对话	
		多模态对话	
		智能图像视频分析	识别能力
			理解能力
			业务拓展性
	服务稳定性		
	无人操作系统	定位导航	
		环境感知	
		规划决策	
		自动控制	
	数据结构化处理与决策	稳定性	
		可用性	
		时效性	
		可解释性	
		可扩展性	

使用成熟度评价参考模型时，应根据不同应用场景构建不同的能力子项评价要素。基于家用智能设备应用场景的能力子项权重和评价方法见附录B。

7.2 能力项描述

7.2.1 AI 基础服务能力

7.2.1.1 存储

存储能力包括：

- 存储管理：数据存储系统的整体管理能力，包括支持的存储容量、数据条目数量等；
- 多样性：支持多种类型数据、多种开发语言的能力；
- 分布式可拓展：支持分布式架构的集群系统业务负载感知、降低运维复杂度的能力；
- 其他：存储服务效率、设备故障预警和报警能力、存储数据恢复能力、接口便捷性和服务可用性等方面。

7.2.1.2 计算

计算能力如下：

- 数据计算：支持超大规模分布式计算、批量计算、流式计算、图计算、机器学习计算和边缘计算等复杂计算任务的能力；
- 集群效率：简单高效的提供大规模集群的部署、调度、扩展和管理服务的能力；

- c) 稳定性：分布式架构部署的稳定性，包括保证故障自动恢复、快速迁移、连续运行时长等能力；
- d) 可扩展性：计算资源扩容、提供扩展接口的能力；
- e) 灵活性：根据业务变化和负载情况灵活调整的能力，以及是否能够提供便捷易用的自动化工具，支持多种框架、控制台、应用程序编程接口和软件开发包的能力；
- f) 统计监控：数据流、数据收发状态的统计监控、报警触发的能力。

7.2.1.3 安全与隐私

安全与隐私保护能力包括：

- a) 协议安全：协议使用可信通道、确保用户的输入不能被第三方窃听的能力；
- b) 数据隐私和算法安全：数据存储中关于用户个人隐私的相关数据采用匿名脱敏存储、考虑算法安全机制，能够通过统计数据反推算法逻辑的能力；
- c) 数据加密：对敏感数据进行加密存储的能力；
- d) 网络隔离：对不同用途和环境的设备进行必要的网络隔离、对不同安全级别的运算环境之间的互通进行权限检查的能力；
- e) 权限管理和鉴权：对数据和服务调用进行鉴权的能力；
- f) 行为审计：对各种操作行为进行审计的能力；
- g) 恢复备份：支撑数据恢复与备份的能力。

7.2.1.4 机器学习

机器学习支撑能力包括：

- a) 全生存周期：支撑包含数据框架（采集、筛选、预标注、标注和脱敏等）、特征探索、模型开发（训练、评估和预测）、模型部署、模型管理、应用发布等全生存周期过程的能力；
- b) 训练性能：支撑具备高性能的运算能力；
- c) 兼容性：兼容其他训练模型代码的能力，包括主流的 API 和训练模型代码等；
- d) 灵活性：灵活分配计算资源的能力，包括训练时灵活分配计算资源的时间颗粒度、灵活的计量计费机制等；
- e) 易用性：操控人工智能服务所使用工具支持可视化的拖拽布局编程服务，组合各种数据源、组件、算法、模型和评估模块的能力；
- f) 算法库：支持提供插件式算子注入，支持多类算子及算法、时空分析、频繁项挖掘等能力。

7.2.2 AI 业务服务能力

7.2.2.1 智能语音

智能语音业务服务能力包括：

- a) 语音识别：噪声场景中唤醒词识别和将语音转换为文本的能力；
- b) 语音合成：将文本转换为语音的能力；
- c) 声纹识别：通过声音判别说话者身份的技术能力。

7.2.2.2 自然语言理解

自然语言理解业务服务能力包括：

- a) 词法、句法分析：支持词法分析、句法分析、语义分析的能力；
- b) 信息抽取：从文本、音频、图像、视频等数据中抽取知识单元的能力；
- c) 信息检索：根据用户需求，查找出所需信息的能力；

- d) 语义推理：根据已有的知识，用规则或模型，扩展相关知识或产生出新知识的能力；
- e) 情感分析：通过对自然语言文本的分析，识别文本作者或文本所描述的个体的情感的能力；
- f) 知识图谱：构建知识图谱并利用知识图谱进行知识计算和推理等方面的能力。

7.2.2.3 智能对话

智能对话业务服务能力包括：

- a) 闲聊：闲聊对话中上下文衔接和提高用户粘性的能力；
- b) 知识型问答：为用户提供知识内容、知识搜索等答疑解惑服务的能力；
- c) 任务导向对话：响应并引导用户的需求，达成任务的能力；
- d) 情感对话：与人对话，结合主题、情景等考虑对人进行情感支持的能力；
- e) 启发式对话：在用户不知道如何做或者能做什么时，给用户提示，引导用户继续对话，考虑信息的利用程度、自主性以及学习的能力；
- f) 多模态对话：在对话中可以同时考虑语音、视频、动作、事件输入，以及表情、动作、语音等多模态输出，考虑模态数量、模态选择策略、学习的能力。

7.2.2.4 智能图像视频分析

智能图像视频分析业务服务能力包括：

- a) 识别能力：支持至少一项识别服务的能力，如人脸识别、物体识别、文字识别等；识别过程中防攻击的能力，例如防呈现攻击；
- b) 理解能力：准确和高效理解图像内容信息的能力；
- c) 业务拓展性：支持不同设备能力情况下的多种不同要求的业务的能力；
- d) 服务稳定性：提供图像视频稳定性的能力，包括容错和负载均衡等。

7.2.2.5 无人操控系统

无人操控系统业务服务能力包括：

- a) 定位和导航：准确获知车辆位置和生成车辆全局行驶路线的能力；
- b) 环境感知：通过多种传感器对车辆周围的环境信息进行感知的能力；
- c) 规划决策：对当前的车辆行为进行规划（例如，速度规划、避障局部路径规划等），并产生相应的决策（例如，跟车、换道、停车等）的能力；
- d) 自动控制：精确、平顺、低延时的执行能力。

7.2.2.6 数据结构化处理与决策

数据结构化处理与决策业务服务能力包括：

- a) 稳定性：数据处理模型对抗噪声和数据攻击的稳定性能力；
- b) 可用性：系统与用户的互协作能力；提供精确的数据分析结果，并给出可用决策建议的能力；在维护和调整时的在线实时服务能力等；
- c) 时效性：提供实时性和变化感知的能力；
- d) 可解释性：数据处理和决策过程中模型的可解释性与流程的透明性；
- e) 可扩展性：跨领域、覆盖多个场景的迁移能力和可扩展度。

附录 A
(规范性附录)
成熟度参考模型评分

服务能力成熟度等级的评分可参考如下公式。其中，各能力主域和能力子域的评分可分别适用相关公式进行评价：

$$f) \text{ 成熟度等级得分} = \sum_k (\text{能力主域 } k \times \text{能力主域权重 } k)$$

其中，能力主域k为构成成熟度评价参考模型的能力主域。

$$g) \text{ 能力主域 } k \text{ 成熟度得分} = \sum_j (\text{能力子域 } j \times \text{能力子域权重 } j)$$

其中，能力子域j为能力主域k下的能力子域。

$$h) \text{ 能力子域 } j \text{ 成熟度得分} = \sum_i (\text{能力子项 } i \times \text{能力子项权重 } i)$$

其中，能力子项i为能力子域j下的能力子项。

人工智能服务能力、某项能力主域、能力子域成熟度分级评分见表A.1，根据得到的相应的成熟度得分，对照相应的服务能力成熟度评级：

表 A.1 人工智能服务能力成熟度分级参考规则

等 级	取值范围	特 征
起始级	1≤得分<1.5	能力主域中各个能力子域的能力子项只达到较低或最低的要求。 人工智能服务只能完成基本的功能，在技术水平、系统性能、可扩展性、安全性等方面只能达到较低或最低的要求。
发展级	1.5≤得分<2.5	能力主域中大多数能力子域的能力子项只达到较低的要求。 人工智能服务能够完成基本的功能，在技术水平、系统性能、可扩展性、安全性等方面中的一部分可以达到较低或中等的要求。
稳健级	2.5≤得分<3.5	能力主域中大多数能力子域的能力子项达到中等的要求。 人工智能服务能够完善的完成基本的功能，在技术水平、系统性能、可扩展性、安全性等方面中的大部分可以达到或超过中等的要求。
优秀级	3.5≤得分<4.5	能力主域中大多数能力子域的能力子项能达到中等或较高的要求。 人工智能服务除了有优秀的功能实现，在技术水平、系统性能、可扩展性、安全性等方面中的大部分可以达到较高的要求。
卓越级	4.5≤得分≤5	能力主域中大多数能力子域的能力子项能达到较高或最高的要求。 人工智能服务除了有卓越的功能实现，在技术水平、系统性能、可扩展性、安全性等方面中的大部分可以达到最高的要求。

附 录 B
(规范性附录)

能力子项权重及成熟度参考模型评分

B.1 能力子项权重

以下为家用智能设备场景、智能医疗诊疗场景、工业设备预测性维护场景分别提供能力主域、能力子域、能力子项权重。各权重以各类因素影响用户体验的重要程度为依据,见表B.1、表B.2、表B.3:

表 B.1 家用智能设备场景能力子项权重

能力主域	能力子域	能力子项	能力子项权重
AI基础服务能力 (权重50%)	存储 (权重25%)	数据管理	25%
		多样性	25%
		分布式可扩展	20%
		其他	30%
	计算 (权重25%)	数据计算	20%
		集群效率	20%
		稳定性	20%
		可扩展性	15%
		灵活性	15%
		统计监控	10%
	安全与隐私 (权重25%)	协议安全	20%
		数据隐私和算法安全	20%
		数据加密	20%
		网络隔离	10%
		权限管理和鉴权	10%
		行为审计	10%
		恢复备份	10%
	机器学习 (权重25%)	全生存周期	20%
		训练性能	15%
兼容性		15%	
灵活性		15%	
易用性		15%	
算法库		20%	
AI业务服务能力 (权重50%)	智能语音 (权重20%)	语音识别	50%
		语音合成	40%
		声纹识别	10%
		词法、句法分析	30%
		信息抽取	15%

表 B.1 家用智能设备场景能力子项权重（续）

能力主域	能力子域	能力子项	能力子项权重
	自然语言理解 (权重20%)	信息检索	15%
		语义推理	15%
		情感分析	15%
		知识图谱	10%
	智能对话 (权重20%)	闲聊	20%
		知识型问答	20%
		任务导向对话	20%
		情感对话	15%
		启发式对话	15%
		多模态对话	10%
	智能图像视频分析 (权重20%)	识别能力	50%
		理解能力	20%
		业务拓展性	15%
		服务稳定性	15%
	数据结构化处理与决策 (权重20%)	稳定性	30%
		可用性	30%
		时效性	15%
可解释性		10%	
可扩展性		15%	

全国团体标准信息平台

表 B.2 智能医疗诊疗场景能力子项权重

能力主域	能力子域	能力子项	能力子项权重
AI基础服务能力 (权重50%)	存储 (权重20%)	数据管理	30%
		多样性	10%
		分布式可扩展	30%
		其他	30%
	计算 (权重30%)	数据计算	20%
		集群效率	20%
		稳定性	30%
		可扩展性	5%
		灵活性	5%
		统计监控	20%
	安全与隐私 (权重30%)	协议安全	10%
		数据隐私和算法安全	30%
		数据加密	20%
		网络隔离	5%
		权限管理和鉴权	5%
		行为审计	10%
		恢复备份	20%
	机器学习 (权重20%)	全生存周期	20%
		训练性能	20%
兼容性		10%	
灵活性		15%	
易用性		25%	
算法库		10%	
AI业务服务能力 (权重50%)	智能语音 (权重20%)	语音识别	60%
		语音合成	40%
		声纹识别	0%
	自然语言理解 (权重20%)	词法、句法分析	30%
		信息抽取	15%
		信息检索	15%
		语义推理	15%
		情感分析	15%
		知识图谱	10%
		智能对话 (权重15%)	闲聊
	知识型问答		20%
	任务导向对话		35%
	情感对话		15%
	启发式对话		10%
	多模态对话		15%

表 B.2 智能医疗诊疗场景能力子项权重（续）

能力主域	能力子域	能力子项	能力子项权重
	智能图像视频分析 (权重25%)	识别能力	60%
		理解能力	10%
		业务拓展性	5%
		服务稳定性	25%
	数据结构化处理与决策 (权重20%)	稳定性	50%
		可用性	15%
		时效性	15%
		可解释性	10%
		可扩展性	10%

全国团体标准信息平台

表 B.3 工业设备预测性维护场景能力子项权重

能力主域	能力子域	能力子项	能力子项权重
AI基础服务能力 (权重50%)	存储 (权重25%)	数据管理	25%
		多样性	25%
		分布式可扩展	20%
		其他	30%
	计算 (权重30%)	数据计算	20%
		集群效率	20%
		稳定性	20%
		可扩展性	15%
		灵活性	15%
		统计监控	10%
	安全与隐私 (权重25%)	协议安全	20%
		数据隐私和算法安全	20%
		数据加密	20%
		网络隔离	10%
		权限管理和鉴权	10%
		行为审计	10%
		恢复备份	10%
	机器学习 (权重20%)	全生存周期	15%
		训练性能	25%
兼容性		15%	
灵活性		10%	
易用性		15%	
算法库		20%	
AI业务服务能力 (权重50%)	自然语言理解 (权重15%)	词法、句法分析	30%
		信息抽取	10%
		信息检索	10%
		语义推理	10%
		情感分析	10%
		知识图谱	30%
	智能图像视频分析 (权重15%)	识别能力	50%
		理解能力	20%
		业务拓展性	15%
		服务稳定性	15%
	数据结构化处理与决策 (权重70%)	稳定性	50%
		可用性	15%
		时效性	15%
		可解释性	10%
可扩展性		10%	

B.2 成熟度参考模型评分

B.2.1 AI 基础服务能力

B.2.1.1 存储

存储能力子域的能力子项成熟度评价分值见表B.4。

表 B.4 存储能力子项成熟度评价分值（1-5 分）

能力子项	评价描述	分值
数据管理	高：支持异构数据的一体化的数据采集技术、EB级分布式存储平台；支持千亿以上数据条目的表格存储系统。	5
	中：支持异构数据的一体化的数据采集技术、较强能力的存储平台。	3
	低：不支持对于不同结构数据的一体化采集技术，受限能力的存储平台。	1
多样性	高：支持多样化的自由存储系统，包括关系数据库、KV存储服务、文件存储服务、结构化存储服务、弹性消息队列等；支持多语言SDK包括Java、Python、PHP、JavaScript、Node.js、Go等语言。	5
	中：支持部分的存储系统以及部分语言SDK。	3
	低：支持特定的存储系统及特定语言SDK。	1
分布式可扩展	高：支持分布式架构的集群系统，系统能够很方便增加或者减少服务器节点个数，以应对业务负载的变化，并且这样的操作是自动化的，减少运维负担。	5
	中：支持分布式架构的集群系统，但无法自动化地根据业务负载调整服务器个数，运维成本适中。	3
	低：不支持分布式架构的集群系统，服务器节点变化对于用户影响较大，运维成本较高。	1
其他	高：提供安全高效的存储服务，自动故障迁移与负载均衡，具备底层资源双可用区支撑，数据强一致保证；系统给用户提简单易懂的库和接口，方便用户使用。	5
	中：支持安全的存储服务，对服务器故障可以快速响应。	3
	低：对服务器故障高度敏感，无法提供安全高效的存储服务。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

B.2.1.2 计算

计算能力子域的能力子项成熟度评价分值见表B.5。

表 B.5 计算能力子项成熟度评价分值（1-5 分）

能力子项	评价描述	分值
数据计算	高：支持基于云计算虚拟化技术的超大规模分布式计算的通用平台；支持批量计算、流式计算、图计算和机器学习计算等复杂任务；支持大规模机器学习平台和异构深度学习平台。	5

表B.5 计算能力子项成熟度评价分值（1-5分）（续）

能力子项	评价描述	分值
	中：支持大规模分布式计算的通用平台；支持批量计算、流式计算、图计算和机器学习计算等复杂任务。	3
	低：不具备大规模的数据计算能力，无法支持复杂度较高的数据计算任务。	1
集群效率	高：能简单高效的提供大规模集群的部署、调度、扩展和管理服务。	5
	中：能较高效的提供较大规模集群的部署、调度、扩展和管理服务，但便捷性有限。	3
	低：能提供初步规模集群的部署、调度、扩展和管理服务，便捷性较低。	1
	稳定性	高：具备底层资源双可用区支撑，保证故障自动恢复、快速迁移，提供7x24高可用服务。
中：保证部分故障能够自动恢复和迁移，提供7x24小时基础服务。		3
低：无法保证故障能够自动恢复和迁移，不能提供7x24小时服务。		1
可扩展性	高：支持低时延（如秒级水平）扩容，提供简洁的预定义扩展接口，用户可以从较低计算规模扩展到较大规模，而运行应用程序的方式并不会因此变得复杂。	5
	中：支持较低时延（多于秒级水平）扩容，提供扩展接口。	3
	低：扩容时延较大，不提供扩展接口。	1
灵活性	高：能够根据业务变化和负载情况，对计算资源进行弹性伸缩，保证计算性能；提供便捷易用的自动化工具，支持多种常用框架和Web控制台、API和SDK等。	5
	中：能够根据业务变化和负载情况之一对计算资源进行调整；提供自动化工具，支持部分常用框架和Web控制台、API和SDK等。	3
	低：不能完全根据业务变化和负载情况对计算资源进行调整；没有自动化工具，只支持个别常用框架、Web控制台、API或SDK。	1
统计监控	高：能够量化整体数据流的数据质量情况，实时统计各模块的数据延迟和收发状态，提供数据流全链路监控，在数据异常和服务异常时能根据规则触发报警。	5
	中：能够量化部分数据流的数据质量情况，能够统计各模块的数据延迟和收发状态，提供数据流部分链路监控，在数据异常和服务异常之一时可以触发报警。	3
	低：无法完善量化数据流的数据质量情况，没有完善统计各模块的数据延迟和收发状态，提供数据流少量链路的监控，只有极个别数据异常和服务异常时有报警机制。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

B.2.1.3 安全与隐私

安全与隐私能力子域的能力子项成熟度评价分值见表B.6。

表B.6 安全与隐私能力子项成熟度评价分值（1-5分）

能力子项	评价描述	分值
协议安全	高：所有的接入协议均基于可信通道进行（例如使用SSL/TLS，对诸如Websocket连接，使用wss等安全协议；对RESTful请求，使用HTTPS安全协议），能够完善的确保用户的输入不会被第三方窃听。	5
	中：部分接入协议通过可信通道进行，能够部分确保用户的输入不被第三方窃听。	3

表B.6 安全与隐私能力子项成熟度评价分值（1-5分）（续）

能力子项	评价描述	分值
	低：只有很少的接入协议通过可信通道进行，不能确保用户输入不被第三方窃听。	1
数据隐私和算法安全	高：数据库中用户隐私相关的数据采用用户id匿名化等机制脱敏存储，用户数据上传时经过用户知情同意；用户画像等隐私数据提取使用时只提供用户特征属性分布，避免单个用户隐私泄露；具有算法隐私保护机制，能够防止逆推得到算法逻辑。	5
	中：数据库中部分的用户隐私相关的数据具备脱敏存储，部分用户数据上传时经过用户知情同意；用户画像等隐私数据提取使用时没有完善的隐私泄露管控；不具备算法隐私保护机制。	3
	低：数据库中无用户隐私相关的数据脱敏存储机制，部分用户数据上传时经过用户知情同意；用户画像等隐私数据提取使用时没有完善的隐私泄露管控；不具备算法隐私保护机制。	1
数据加密	高：对重要的敏感数据有足够可靠的加密机制完成加密存储，加密对开发人员和数据库管理人员均有效。	5
	中：对重要的敏感数据有普通的加密机制完成加密存储，加密对部分的开发人员和数据库管理人员有效。	3
	低：对重要的敏感数据只有简单的加密机制，开发人员或数据库管理人员可以绕过加密机制读取数据。	1
网络隔离	高：对不同用途和环境的机器，进行严格的网络隔离；线上生产环境和线下测试开发环境不能互相访问。	5
	中：对不同用途和环境的机器，有网络隔离机制；线上生产环境和线下测试开发环境有简单访问的控制。	3
	低：对不同用途和环境的机器，没有完善的网络隔离机制；线上生产环境和线下测试开发环境没有访问控制。	1
权限管理和鉴权	高：对服务间调用建立了基于密码学和权限列表的访问控制，避免通过 API 窃取用户数据；高安全级别的运算环境和低安全级别的运算环境只有通过严格的权限检查后才可互通。	5
	中：对服务间调用建立了权限列表的访问控制，但没有完善的加密机制；不同安全级别的运算环境之间的互通有简单的权限检查机制。	3
	低：对服务间调用没有完善的权限列表的访问控制；不同安全级别的运算环境之间的互通没有权限检查机制。	1
行为审计	高：对所有登录后的操作进行记录，在必要的时间范围内可以进行检查。	5
	中：对部分登录后的操作进行记录，在较短的时间范围内可以进行检查。	3
	低：没有完善的操作记录机制，无法在一定时间范围内进行检查。	1
恢复备份	高：恢复数据、服务的完整性高（大于等于99%），允许数据、服务的中断时长低（小于5分钟）。	5
	中：恢复数据、服务的完整性一般（大于等于80%，小于99%），允许数据、服务的中断时长一般（大于等于5分钟，小于30分钟）。	3
	低：恢复数据、服务的完整性低（小于80%），允许数据、服务的中断时长高（大于30分钟）。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

B.2.1.4 机器学习

机器学习支撑能力子域的能力子项成熟度评价分值见表B.7。

表 B.7 机器学习集群能力子项成熟度评价分值（1-5 分）

能力子项	评价描述	分值
全生存周期	高：支持包含数据框架（采集、筛选、预标注、标注、脱敏等）、特征探索、模型开发（训练、评估和预测）、模型部署、模型管理、应用发布等全生存周期过程。	5
	中：支持以上大部分生命周期过程。	3
	低：只支持以上少量的生命周期过程。	1
训练性能	高：支持超高性能GPU运算（例如NVIDIA Tesla K80等），具备支持较高内存以上的单机运算能力；支持数据并行和模型并行、单机多卡和多机多卡的分布式训练；同时具备CPU服务器集群、虚拟机集群计算能力。	5
	中：支持较高性能GPU运算，具备支持普通内存以上的单机运算能力；部分支持数据并行和模型并行、单机多卡和多机多卡的分布式训练；部分具备CPU服务器集群、虚拟机集群计算能力。	3
	低：不支持高性能GPU运算能力，只有简单的并行计算能力；不能完善的具备CPU服务器集群、虚拟机集群计算能力。	1
兼容性	高：支持常见的深度学习框架的标准API（例如TensorFlow等），兼容常见的samples代码（例如Google CloudML），相同模型代码可在不同云平台上训练。	5
	中：支持部分的深度学习框架的标准API，兼容部分samples代码。	3
	低：只支持自身设备的深度学习框架的API，不兼容其他samples代码。	1
灵活性	高：支持按需申请和分配CPU、内存和GPU资源。	5
	中：在部分维度上支持按需申请和分配CPU、内存和GPU资源。	3
	低：不支持按需申请和分配CPU、内存和GPU资源。	1
易用性	高：支持简单易用的命令行工具（例如可在Linux/Mac/Windows操作系统或者Docker中运行），可以通过API、SDK或者Web控制台使用云深度学习服务；支持易用的方式提交云端训练，支持多个主流深度学习框架和超参数自动调优等功能；支持可视化的拖拽布局编程服务，组合各种数据源、组件、算法、模型和评估模块。	5
	中：支持部分易用的工具，部分可以通过API、SDK或者Web控制台使用云深度学习服务；支持云端训练；支持部分功能的可视化。	3
	低：所使用的工具易用性不高，无法通过API、SDK或者Web控制台使用云深度学习服务；不支持云端训练；不支持可视化。	1
算法库	高：支持提供插件式算子注入能力；支持70种类型以上的算子及算法，支持时空分析，支持频繁项挖掘。	5
	中：支持少量的以上算法能力。	3
	低：只支持个别的或不支持以上的算法能力。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

B.2.2 AI 业务服务能力

B.2.2.1 智能语音

智能语音能力子域的能力子项成熟度评价分值见表B.8。

表 B.8 智能语音能力子项成熟度评价分值（1-5 分）

能力子项	评价描述	分值
语音识别	高：在不同信噪比状态下，识别率很高（大于90%）而且误识率很低。	5
	中：在不同信噪比状态下，识别率较高（小于等于90%，高于等于80%）而且误识率较低。	3
	低：在不同信噪比状态下，识别率较低（小于80%）或者误识率较高。	1
语音合成	高：分词准确，句读流畅，用户可以准确理解，与自然人类发声没有显著区别。	5
	中：分词较准确，句读较为流畅，用户基本可以理解，与自然人类发声有一定区别。	3
	低：分词不准确，句读不流畅，影响用户理解，与自然人类发声有较大区别。	1
声纹识别	高：错误接受率小于等于0.25%，错误拒绝率小于等于1.5%。	5
	中：错误接受率小于等于0.5%，错误拒绝率小于等于3%。	3
	低：错误接受率小于等于0.75%，错误拒绝率小于等于5%。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

B.2.2.2 自然语言理解

自然语言理解能力子域的能力子项成熟度评价分值见表B.9。

表 B.9 自然语言理解能力子项成熟度评价分值（1-5 分）

能力子项	评价描述	分值
词法、句法分析	高：字、词、句识别和语义理解错误率很低（小于5%）、召回率高(或称命中率不低于90%)，可以根据上下文理解复杂句子，响应时间符合自然人类问答习惯，完全不影响AI语音服务功能和语音交互用户体验。	5
	中：字、词、句识别和语义理解错误率较低（小于10%）、召回率较高(或称命中率不低于75%)，可以理解复杂句子，对AI语音服务功能和语音交互用户体验有少量的影响。	3
	低：字、词、句识别和语义理解错误率较高（大于等于20%），对AI语音服务功能和语音交互用户体验有明显影响。	1
信息抽取	高：能提取关键词，命中实体及实体间的显示语义关系，结合上下文提取实体间的隐含语义关系、背景知识、常识知识。	5
	中：能提取关键词，命中实体及实体间的显示语义关系，结合上下文提取实体间的隐含语义关系。	3
	低：能提取关键词，命中实体及实体间的显示语义关系。	1
信息检索	高：满足所有信息需求，根据用户意图对结果做准确排序。	5
	中：满足基本和重要信息需求，能够对结果做简单排序。	3
	低：满足基本信息需求。	1

表 B.9 自然语言理解能力子项成熟度评价分值（1-5 分）（续）

能力子项	评价描述	分值
语义推理	高：可以实现通用的逻辑推理，并拥有常识，准确率大于等于90%，召回率（或称命中率）不低于80%。	5
	中：可以实现特定领域的逻辑推理，准确率大于等于90%，召回率（或称命中率）不低于50%。	3
	低：不能进行推理。	1
情感分析	高：可以准确区分10种以上（包括但不限于快乐、着急、愤怒、热情、鄙视、哀伤等）不同的情绪，每一种情绪区分至少3个级别（准确率大于90%）。	5
	中：可以准确区分（准确率大于90%）正面、负面或中性的情绪。	3
	低：无法相对准确地区分正面、负面或中性的情绪。	1
知识图谱	高：可以支持通用领域和专用领域知识图谱的构建，支持利用知识图谱进行跨领域的知识计算和推理，支持多个语种。	5
	中：支持专用领域知识图谱的构建，支持利用知识图谱进行知识计算和推理，支持多个语种。	3
	低：支持专用领域知识图谱的构建，支持利用知识图谱进行知识计算和推理，支持单个语种。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

B.2.2.3 智能对话

智能对话能力子域的能力子项成熟度评价分值见表B.10。

表 B.10 智能对话能力子项成熟度评价分值（1-5 分）

能力子项	评价描述	分值
闲聊	高：平均对话单轮次数高，上下文衔接很流畅。	5
	中：平均对话单轮次数中，上下文衔接很较流畅。	3
	低：平均对话单轮次数低，上下文衔接不流畅。	1
知识型问答	高：能够在限定和公开的知识库中根据关键词查询答案；根据标注数据，学会使用明确的语法来表示用户的问题，并自动转换为不同知识库支持的查询语句；根据已有的知识图谱识别用户意图，给出相应回复，能够根据交互及世界知识自动进行知识图谱扩充和完善；根据用户问题和答案，自动学会使用语法表示用户问题及执行查询；根据已有会话，自动识别其中的上下文，以及如何相应的应答策略；依靠少量标注数据和用户反馈，能逐步学会回答新领域里的问题。	5
	中：能够在限定和公开的知识库中根据关键词查询答案；根据已有的知识图谱识别用户意图，给出相应回复；有利用上下文信息的能力，依靠少量标注数据和用户反馈，能逐步学会回答新领域里的问题。	3
	低：只能在限定领域的知识库中根据关键词查找答案。	1

表 B.10 智能对话能力子项成熟度评价分值（1-5 分）（续）

能力子项	评价描述	分值
任务导向对话	高：能够完成单句指令任务和多轮交互型任务。在多轮交互型任务上，支持带有固定服务目标的对话引导；可以在用户目标及其目标间寻找最优目标组合，并以此组合目标为基础，支持基于用户个性化数据的带目标的任务型对话，并能自主选择最佳的交互模态，可参数化设定性能指标。	5
	中：能够完成单句指令任务和多轮交互型任务。在多轮交互型任务上，支持带有固定服务目标的对话引导。	3
	低：能够完成单句指令任务和多轮交互型任务。	1
情感对话	高：在多轮对话中，理解并引导用户情绪，支持视频等多模态交互；在与用户会话过程中，仅根据用户的反馈，自动学会应对不同的用户情绪、主题和场景；可以发掘用户面临的问题或挑战，并提出新的观点，并引导用户接受新观点，可参数化设定性能指标。	5
	中：在多轮对话中，能够识别主人情感，能结合当前主题和情景，给予合适应答以达成情感支持目的；可以根据标注数据，学会应对不同的用户情绪、主题和场景。	3
	低：单轮对话中，能够识别主人的正向、负向、中性的情感，并采用相应情感应答。	1
启发式对话	高：根据标注好的数据，学会新的触发条件和引导方式；根据用户画像、具有相近用户画像的用户交互历史，预测用户最可能感兴趣的对话内容；可以从已有的交互记录以及用户对历史引导的反馈数据中，自动学会何时以及怎样引导用户最好；可以根据用户对引导的反馈，自主调整引导策略；能够参数化设定性能指标。	5
	中：根据标注好的数据，学会新的触发条件和引导方式；根据用户画像、具有相近用户画像的用户交互历史，预测用户最可能感兴趣的对话内容。	3
	低：根据一定的策略和推荐规则给出问题，能自然过渡到预先设定的目标话题。	1
多模态对话	高：满足所有信息需求，能根据用户意图对结果进行准确排序，重要结果置顶。	5
	中：满足所有信息需求，能对结果进行简单排序。	3
	低：满足基本和重要信息需求。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

B.2.2.4 智能图像视频分析

智能图像视频分析子域的能力子项成熟度评价分值见表B.11。

表 B.11 智能图像视频分析能力子项成熟度评价分值（1-5 分）

能力子项	评价描述	分值
识别能力	高：至少支持一项识别服务如人脸识别、物体识别、文字识别等，识别准确率较高；支持防2种以上呈现攻击形式（3D、2D、静态图像、动态图像、面具、头模）等的的能力，防呈现攻击能力较高。	5
	中：至少支持一项识别服务如人脸识别、物体识别、文字识别等，识别准确率一般；支持1到2种呈现攻击形式能力，防呈现攻击能力一般。	3

表B.11 智能图像视频分析能力子项成熟度评价分值（1-5分）（续）

能力子项	评价描述	分值
	低：只支持特定的识别服务，识别准确率较低；防呈现攻击能力弱或者不具备，比较容易进行攻击。	1
理解能力	高：能够良好解读图像的内容信息，识别正确率高于等于90%，解读效率高、用时短，完全不影响用户体验。	5
	中：能够完成解读图像内容信息的功能，识别正确率较高高于等于80%，低于90%，解读效率和用时适中，只少量影响用户体验。	3
	低：能够完成解读图像内容信息的功能，识别正确率明显影响用户体验（低于80%），解读效率低，用时较长。	1
业务拓展性	高：支持不同设备能力情况下的多种不同要求的业务，支持基于云平台 and 近端的视觉服务。	5
	中：支持多种不同要求的业务，仅支持基于近端的视觉服务。	3
	低：仅支持对于时延要求不敏感的业务，仅支持基于近端的视觉服务。	1
服务稳定性	高：提供由多台高性能服务器组成的集群的服务后台，拥有较高的容错和负载均衡能力，为客户提供稳定高效的图像视频服务。	5
	中：服务处理能力强，不能够提供持续稳定的服务。	3
	低：服务处理能力较弱，不能够提供持续稳定的服务。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

B.2.2.5 无人操控系统

无人操控系统的能力子项成熟度评价分值见表B.12。

表B.12 无人操控系统能力子项成熟度评价分值（1-5分）

能力子项	评价描述	分值
定位导航	高：高复杂环境下，定位精度高，位置更新速度快，位置信息无丢失；具备全程导航、局部导航、避障等能力；满足完全自动驾驶车辆告诉行驶的要求。	5
	中：高复杂环境下，定位精度高，位置更新速度快；特定环境下，全自动无人操控车辆具备局部导航、避障等能力；满足辅助自动驾驶车辆的要求。	3
	低：一般环境下，定位精度满足人工驾驶的需要；特定环境下，自动驾驶车辆具备辅助驾驶人员导航及避障的能力。	1
环境感知	高：复杂情况下，能够完全感知交通信号标识、路面情况信息及周围环境等信息。	5
	中：特定情况下，能够部分感知交通信号标识、路面情况信息等信息，准确度高。	3
	低：特定情况下，能够部分感知交通信号标识、路面情况信息等信息，而且准确度不高，不具备自动驾驶的能力。	1
规划决策	高：能根据路网信息、获取的交通环境信息和自身行驶状态，自主全程决定车辆的驾驶状态。	5
	中：能根据路网信息、获取的交通环境信息和自身行驶状态，在特定情况下，具备自主决定车辆的驾驶状态。	3

表 B. 12 无人操控系统能力子项成熟度评价分值（1-5 分）（续）

能力子项	评价描述	分值
	低：能根据路网信息、获取的交通环境信息和自身行驶状态，辅助驾驶人员操控车辆。	1
自动控制	高：复杂环境下，能够根据规划的行驶轨迹和速度，以及当前的自身行驶状态，产生对车辆的控制命令，车辆运行轨迹与规划轨迹重合度高。	5
	中：特定环境下，能够根据规划的行驶轨迹和速度，以及当前的自身行驶状态，产生对车辆的控制命令，车辆运行轨迹与规划轨迹基本重合。	3
	低：特定情况下，能够根据规划的行驶轨迹和速度，以及当前的自身行驶状态，产生对车辆的控制命令，但不能使车辆运行轨迹与规划轨迹基本重合。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

B. 2. 2. 6 数据结构化处理与决策

数据结构化处理与决策的能力子项成熟度评价分值见表B. 13。

表 B. 13 智能数据处理与决策能力子项成熟度评价分值（1-5 分）

能力子项	评价描述	分值
稳定性	高：系统能完全管理且容易配置；针对噪声污染、对抗数据攻击等情形，所提供的智能分析业务稳定，输出结果一致。	5
	中：基本可管理，配置有一定难度，仅能针对噪声污染提供有限的业务稳定性。	3
	低：完全不可管理，配置困难，所提供的智能分析业务仅能在较为理想的条件下提供业务，扰动下输出结果可能不一致。	1
可用性	高：能够提供高准确率（大于等于95%）、低错误率（小于等于5%）、高灵敏度（大于等于95%）、高特异性（大于等于95%）的分类精度；决策科学合理，可行性很高，易于接受；支持用户在数据处理与决策的全过程中实时交互，同时系统可为每位用户建立自适应的协同模型；服务可在不暂停服务的同时开展完全的系统维护和算法、模型等模块更新。	5
	中：能够提供较高准确率（大于等于80%，小于95%）、较低错误率（大于5%，小于等于20%）、较高灵敏度（大于等于80%，小于95%）、较高特异性（大于等于80%，小于95%）的分类精度，错误分类可能降低生产力；决策基本合理、可行，可以接受；在大部分环节中（大于60%，小于100%）支持用户在数据处理与决策中与系统交互，系统可为部分用户建立自适应的协同模型；超过50%但低于100%的系统维护和算法、模型等模块更新操作时，无需暂停服务。	3
	低：仅能提供较高准确率（大于等于90%，小于95%），但可能出现灵敏度（小于80%）或特异性较低情形（小于80%），影响用户对分类结果的使用；决策不合理，不可行，无效，无法接受；仅在较少环节（小于20%）支持用户在数据处理与决策中与系统交互，系统不支持为用户建立自适应协同模型；系统维护和算法、模型等模块更新时需暂停服务。	1

表B.13 智能数据处理与决策能力子项成熟度评价分值（1-5分）（续）

能力子项	评价描述	分值
时效性	高：快速反应，在一定的时间限制和要求内能够快速处理并给出结果，并能感知变化，如果方案已经过时，可根据新的情况及时作出调整。	5
	中：及时反应，在一定的时间限制和要求内能够及时处理并给出结果，保持实施期间的相对稳定性。	3
	低：延迟反应，在处理和决策时拖延时间，决策的效益低，甚至会因问题的性质发生根本变化，而使决策无效。	1
可解释性	高：数据处理流程和分析结果具有完全的可解释性和可预测性，处理和决策过程可理解，业务可阐释性高，可呈现为供有资质第三方审查的文本（超过90%），不存在黑盒。	5
	中：数据处理流程存在无法预判的随机分支，处理和决策过程可基本理解，业务可基本阐释，决策模型方法基本可解释说明（超过50%），但更换同领域其他数据集时，90%以上分析结果可预测。	3
	低：数据处理流程可解释性差，处理和决策过程难理解，业务无法阐释，决策行为很难甚至无法解释说明（低于50%）。	1
可扩展性	高：面向一个领域的服务易于迁移至另一业务领域，同时能够综合多种智能处理和决策方法，可处理多个场景的知识，并做出有效决策。	5
	中：仅能针对某些共性领域开展稳定可靠的服务，适用某些场景的知识处理，针对某些特定的问题可做出有效决策。	3
	低：仅针对某一特定领域提供服务或只适用于单个特定场景。	1
注：以上能力未达到最低评分要求或不具备该能力项功能的应评为0分。		

中国电子工业标准化技术协会（CESA）是全国电子信息产业标准化组织和标准化工作者自愿组成的社会团体。广泛联系全国电子信息产业标准化机构和标准化工作者，协助政府部门搞好电子信息产业标准化工作，开拓信息技术领域的标准化工作是中国电子工业标准化技术协会的主要工作内容之一。中国境内从事科研开发、制造、营销和服务的企事业单位、高等院校、社会组织和个人均可随时向中国电子工业标准化技术协会团体标准工作部提出团体标准项目建议。

中国电子工业标准化技术协会标准按照《电子工业标准化技术协会协会团体标准管理办法》进行制定和管理。

在本标准实施过程中，如发现需要修改或补充之处，请将意见和有关资料寄至中国电子工业标准化技术协会，以便修订时参考。

全国团体标准信息平台

本标准版权归中国电子工业标准化技术协会所有。

中国电子工业标准化技术协会地址：北京市海淀区万寿路27号

电话：010 - 64102952

电子邮箱：standards@cesa.cn

网址：www.cesa.cn
