

TC260

TC260-003

生成式人工智能 服务安全基本要求

Basic security requirements for generative artificial intelligence service

2024-02-29 发布

全国 安全标准化技术委员会发布

次

	II
1	1
2	1
3	1
4	1
5	2
5.1	2
5.2	2
5.3	3
6	3
7	4
8	5
8.1	5
8.2	5
8.3	6
8.4	6
9	6
9.1	6
9.2	7
9.3	7
9.4	7
A	8
	10

前



1

2 性引 文件

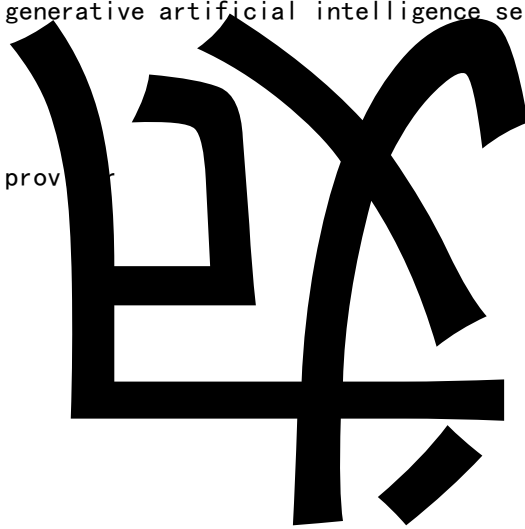
3 术P 和定义

3.1

成式人工智 服务 generative artificial intelligence service

3.2

服务提供 service prov



5 料安全 求

5.1 料来源安全 求

注1:

注2:

注3:

5.2 料内容安全 求

5.3 料标注安全 求

6 模型安全 求

注：

7 安全措施 求

注：

8 其他 求

8.1 关 库

8.2 成内容测 库

8.3 拒 测 库

8.4 分 模 型

9 安全 估 求

9.1 估方法

注1:

注2:

注3:

9.2 料安全 估

9.3 成内容安全 估

9.4 拒 估

录 A

(性)

料及 成内容 主 安全

A. 1 包含 反 会主义核心价值 内容

A. 2 包含歧 性内容

A. 3 商业 法

A. 4 侵犯他人合法权

A.5 无法满 特定服务 型 安全 求

参 文 献