

TC260-PG-20258A

网络安全标准实践指南

——人工智能生成合成内容标识方法 文件
元数据隐式标识 文本文件

(V1.0-202508)



全国网络安全标准化技术委员会秘书处

2025 年 08 月

本文档可从以下网址获得：

www.tc260.org.cn/



全国网络安全标准化技术委员会
National Technical Committee 260 on Cybersecurity of SAC



前 言

《网络安全标准实践指南》（以下简称《实践指南》）是全国网络安全标准化技术委员会（以下简称“网安标委”）秘书处组织制定和发布的标准相关技术文件，旨在围绕网络安全法律法规政策、标准、网络安全热点和事件等主题，宣传网络安全相关标准及知识，提供标准化实践指引。

本文件起草单位：北京金山办公软件股份有限公司、中国电子技术标准化研究院、国家计算机网络应急技术处理协调中心、浙江大学、杭州高新区（滨江）区块链与数据安全研究院、珠海金山办公软件有限公司、北京快手科技有限公司、阿里云计算有限公司、武汉金山办公软件有限公司、中央网信办数据与技术保障中心、国家计算机网络应急技术处理协调中心山西分中心、北京抖音信息服务有限公司、腾讯音乐娱乐（深圳）有限公司、小米科技有限责任公司、荣耀终端股份有限公司、北京大麦地信息技术有限公司、深圳市爱思软件技术有限公司、维沃移动通信有限公司。

本文件主要起草人：安红云、刘丹、任奎、许晓耕、贺敏、张树玲、姚冬、王少康、秦湛、王志伟、郝春亮、张立尧、张博然、孙培尧、董琳、吴峰、朱云峰、朱军行、何文甲、牛文杰、刘卓月、杨敏、陈一帆、苏永新、何求胜、高思宁、落红卫、谷晨、贺凯、杜蕾、李茹、李伟峰、罗雄清、朱柄丞、张莹、李辰淑、孙勇、史骥、徐曼。



声 明

本《实践指南》版权属于网安标委秘书处，未经秘书处书面授权，不得以任何方式抄袭、翻译《实践指南》的任何部分。凡转载或引用本《实践指南》的观点、数据，请注明“来源：全国网络安全标准化技术委员会秘书处”。



全国网络安全标准化技术委员会
National Technical Committee 260 on Cybersecurity of SAC



摘 要

为落实《人工智能生成合成内容标识办法》，根据强制性国家标准 GB 45438—2025《网络安全技术 人工智能生成合成内容标识方法》的要求，本文件提供了人工智能生成合成的文本文件元数据隐式标识方法，包括针对不同文件格式的具体方案，指导人工智能生成合成内容服务提供者和网络信息内容传播服务提供者开展人工智能生成合成的文本文件元数据隐式标识活动。





目 录

1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	2
5 概述	2
5.1 文本文件格式	2
5.2 方法概述	3
6 文本文件的元数据隐式标识方案	3
6.1 基于 OOXML 的文本文件	3
6.2 基于 UOF 的文本文件	3
6.3 基于 PDF 的文本文件	4
6.4 基于 OFD 的文本文件	4
6.5 基于 Markdown 的文本文件	5
6.6 基于 xmind 的文本文件	5
6.7 基于 posm 和 posf 的文本文件	6
附录 A （资料性）读写文本文件元数据隐式标识伪代码	7
附录 B （资料性）文本文件元数据隐式标识示例	9
参考文献	15





1 范围

本文件给出了人工智能生成合成文本内容的文件元数据隐式标识方法，包括针对不同文本文件格式的具体方案。

本文件适用于指导人工智能生成合成文本内容服务提供者和网络信息内容传播服务提供者开展文件元数据隐式标识活动。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；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 45438—2025 网络安全技术 人工智能生成合成内容标识方法

3 术语和定义

下列术语和定义适用于本文件。

3.1 人工智能生成合成文本内容 **text content generated by artificial intelligence**

利用人工智能技术生成、合成的文本信息。

3.2 人工智能生成合成文本内容服务提供者 **artificial intelligence text content generation service provider**

生成合成文本服务提供者

利用人工智能技术（包括通过提供可编程接口等方式）向公众提



供生成合成文本服务的组织或个人。

3.3 网络信息文本内容传播服务提供者 internet information text content propagation service provider

文本内容传播服务提供者

提供网络信息文本内容传播服务的网络信息服务提供者。

4 缩略语

下列缩略语适用于本文件。

OFD: 开放版式文档 (Open Fixed layout Document)

OOXML: 开放办公文档格式 (Office Open XML)

PDF: 便携式文档格式 (Portable Document Format)

UOF: 标文通 (Uniform Office Document Format)

5 概述

5.1 文本文件格式

文本文件格式一般包括:

- a) 基于 OOXML 规范的文本文件, 如 docx、pptx、xlsx、dotx、xltx、potx;
- b) 基于 UOF 规范的文本文件, 如 uot、uott、uos、uost、uop、uopt;
- c) 基于 PDF 规范的文本文件, 如 pdf;
- d) 基于 OFD 规范的文本文件, 如 ofd;
- e) 基于 Markdown 规范的文本文件, 如 md;
- f) 基于 xmind 规范的文本文件, 如 xmind;



g) 基于 posm、posf 规范的文本文件，如 posm、posf。

5.2 方法概述

针对本文件 5.1 所列各类格式的文本文件，本文件给出了元数据隐式标识方案，其他文件格式可参考处理。

6 文本文件的元数据隐式标识方案

6.1 基于 OOXML 的文本文件

基于 OOXML 的文本文件，方案如下：

a) 存储路径为 docProps/custom.xml 文件中的元素：

Properties/property;

b) 写入方式为：

1) 新建元素 property，其属性 name 为“AIGC”；

2) 子元素 lpwstr 的值采用 json 字符串表示，字符串如下：

```
{"Label": "value1", "ContentProducer": "value2", "ProduceID": "value3", "ReservedCode1": "value4", "ContentPropagator": "value5", "PropagateID": "value6", "ReservedCode2": "value7"}
```

基于 OOXML 的文本文件元数据读写方式参见附录 A，实现示例参见附录 B。

6.2 基于 UOF 的文本文件

基于 UOF 的文本文件，方案如下：

a) 存储位置为 _meta/meta.xml 中的元素：自定义元数据集_520F/



自定义元数据_5210;

b) 写入方式为:

1) 新建元素“自定义元数据_5210”，其属性“名称_5211”为
“AIGC”;

2) “自定义元数据_5210”的值:

```
{"Label": "value1", "ContentProducer": "value2", "ProduceID": "value3", "ReservedCode1": "value4", "ContentPropagator": "value5", "PropagateID": "value6", "ReservedCode2": "value7"}.
```

基于 UOF 的文本文件元数据读写方式参见附录 A，实现示例参见附录 B。

6.3 基于 PDF 的文本文件

基于 PDF 的文本文件，方案如下:

a) 存储位置为 Document Information Dictionary;

b) 写入方式为:

```
/AIGC({ "Label": "value1", "ContentProducer": "value2", "ProduceID": "value3", "ReservedCode1": "value4", "ContentPropagator": "value5", "PropagateID": "value6", "ReservedCode2": "value7" }).
```

基于 PDF 的实现示例参见附录 B。

6.4 基于 OFD 的文本文件

基于 OFD 的文本文件，方案如下:

a) 存储位置为 OFD/DocBody/DocIno 文件中的元素:



CustomDatas

b) 写入方式为:

1) 新建元素 CustomData, 其属性 name 为“AIGC”;

2) CustomData 的值:

```
{"Label":"value1","ContentProducer":"value2","ProduceID":"value3","ReservedCode1":"value4","ContentPropagator":"value5","PropagateID":"value6","ReservedCode2":"value7"}。
```

基于 OFD 的文本文件元数据读写方式参见附录 A, 实现示例参见附录 B。

6.5 基于 Markdown 的文本文件

基于 Markdown 的文本文件, 方案如下:

a) 存储位置为嵌入在 Markdown 文件头部;

b) 写入方式为见图 1:

```
---  
AIGC:  
  Label: 'value 1'  
  ContentProducer: 'value 2'  
  ProduceID: 'value 3'  
  ReservedCode1: 'value 4'  
  ContentPropagator: 'value 5'  
  PropagateID: 'value 6'  
  ReservedCode2: 'value 7'  
---
```

图 1 Markdown 元数据标识

6.6 基于 xmind 的文本文件

基于 xmind 的文本文件, 方案如下:

a) 存储位置为嵌入在 metadata.json 文件下新建 AIGC 节点, 其值为 json 字符串;



b) 写入方式如下:

```
{"Label":"value1","ContentProducer":"value2","ProduceID":"value3","ReservedCode1":"value4","ContentPropagator":"value5","PropagateID":"value6","ReservedCode2":"value7"}。
```

6.7 基于 posm 和 posf 的文本文件

基于 posm 和 posf 的文本文件, 方案如下:

a) 存储位置为嵌入在 meta.json 文件下新建 AIGC 节点, 其值为 json 字符串;

b) 写入方式如下:

```
{"Label":"value1","ContentProducer":"value2","ProduceID":"value3","ReservedCode1":"value4","ContentPropagator":"value5","PropagateID":"value6","ReservedCode2":"value7"}。
```





附录 A

(资料性)

读写文本文件元数据隐式标识伪代码

为了直观表达元数据隐式标识的算法和操作步骤，元数据隐式标识伪代码采用自然语言和编程语言混合的算法描述方法，表达文本文件元数据隐式标识算法。

以下伪代码示例是基于 OOXML、UOF、OFD 的文本文件元数据属性读写：

a) 读取元数据属性：

函数 读取元数据(文档路径)：

打开文档

创建标准元数据字典

创建自定义属性字典

// 读取自定义属性

打开 ZIP 文档：

如果存在自定义属性文件：

 读取 custom.xml 文件内容

 解析 XML 根节点

// 遍历所有属性节点

对每个属性节点：

 获取属性名称

 获取属性值

 将属性名和值添加到自定义属性字典

返回 合并后的标准元数据和自定义属性字典

b) 写入元数据属性

函数 写入元数据(文档路径，属性名，属性值)：

将属性值转换为字符串

创建临时文件路径

// 复制原文件内容到临时文件(排除 custom.xml)

以只读模式打开原文档 ZIP 包：

 以写入模式打开临时 ZIP 包：



遍历原文档所有文件:

如果不是 custom.xml 文件:

复制文件到临时 ZIP 包

// 处理自定义属性 XML

以追加模式打开临时 ZIP 包:

尝试:

读取现有 custom.xml 内容

解析 XML 根节点

否则:

创建新的 XML 根节点 (Properties)

// 查找或创建属性节点

查找属性名匹配的节点

如果未找到:

创建新的属性节点

设置属性 ID 和格式 ID

创建值节点并设置值

否则:

更新现有节点的值

// 将修改后的 XML 写回 ZIP 包

将 XML 根节点写入 custom.xml 文件



附录 B

(资料性)

文本文件元数据隐式标识示例

B.1 docx 文件

docx 文件元数据标识结果见图 2, docx 文件元数据标识展示界面见图 3。



图 2 docx 文件元数据标识结果示例



图 3 docx 文件元数据展示界面示例



B.2 xlsx 文件

xlsx 文件元数据标识结果见图 4，xlsx 文件元数据标识展示界面见图 5。

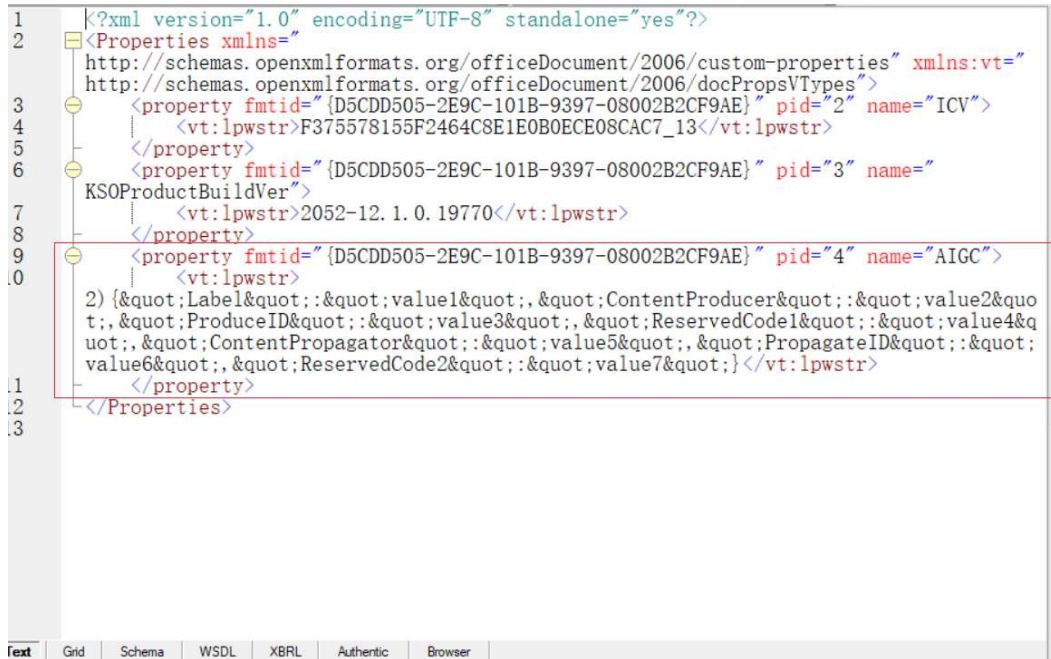


图 4 xlsx 文件元数据标识结果示例

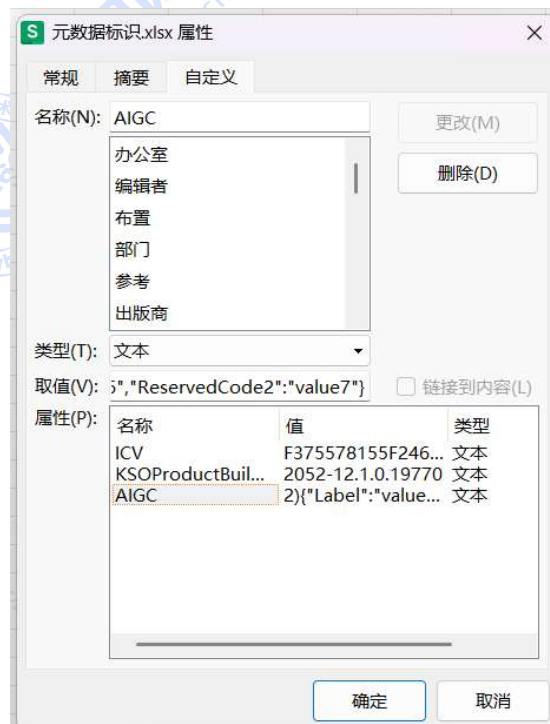


图 5 xlsx 文件元数据标识展示界面示例



B.3 pptx 文件

pptx 文件元数据标识结果见图 6，pptx 文件元数据标识展示界面见图 7。

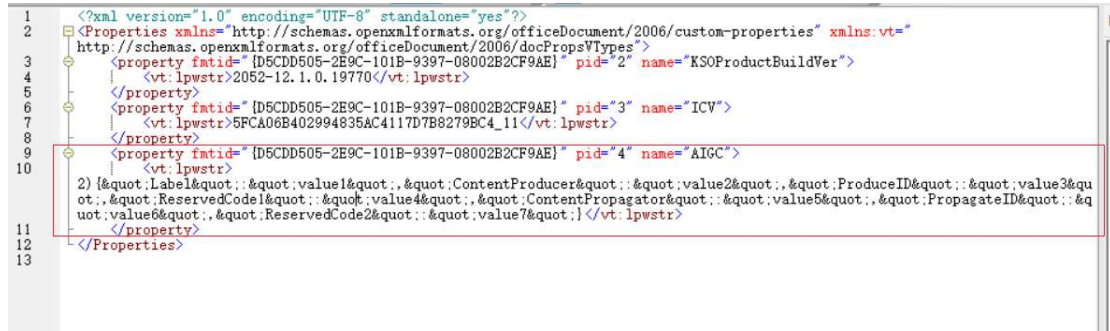


图 6 pptx 文件元数据标识结果示例



图 7 pptx 文件元数据标识展示界面示例

B.4 uot 文件

uot 文件元数据标识结果见图 8，uot 文件元数据标识展示界面见图 9。



图 8 uot 文件元数据标识结果示例



图 9 uot 文件元数据标识展示界面示例

B.5 pdf 文件

pdf 文件元数据标识结果见图 10。

```
172 0 obj
<< /ModDate (D:20250407152043+08'00') /Trapped(False) /Producer()
/AIGC({"Label":"value1","ContentProducer":"value2","ProduceID":"value3","ReservedCode1":"v
alue4","ContentPropagator":"value5","PropagateID":"value6","ReservedCode2":"value7"})
/SourceModified(D:20250217152418+08'00') /Comments() /Title(标准文档) /Keywords()
/Author(wps)/Creator(WPS文字) /CreationDate (D:20250217152418+08'00') /Company()>>
endobj
```

图 10 pdf 文件元数据标识结果示例



B.6 ofd 文件

ofd 文件元数据标识结果见图 11，ofd 文件元数据标识展示界面见图 12。



图 11 ofd 文件元数据标识结果示例

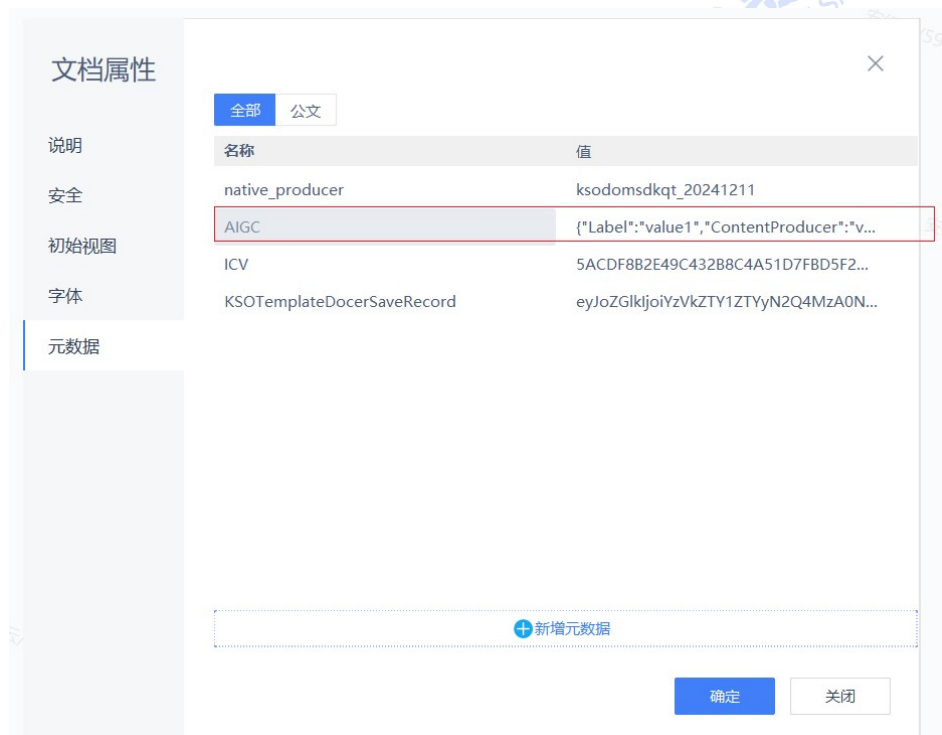


图 12 OFD 文件元数据标识展示界面示例

B.7 xmind 文件

xmind 文件元数据标识结果见图 13。



图 13 xmind 文件元数据标识结果示例

B.8 posm 和 posf 文件

posm 和 posf 文件元数据标识结果见图 14。

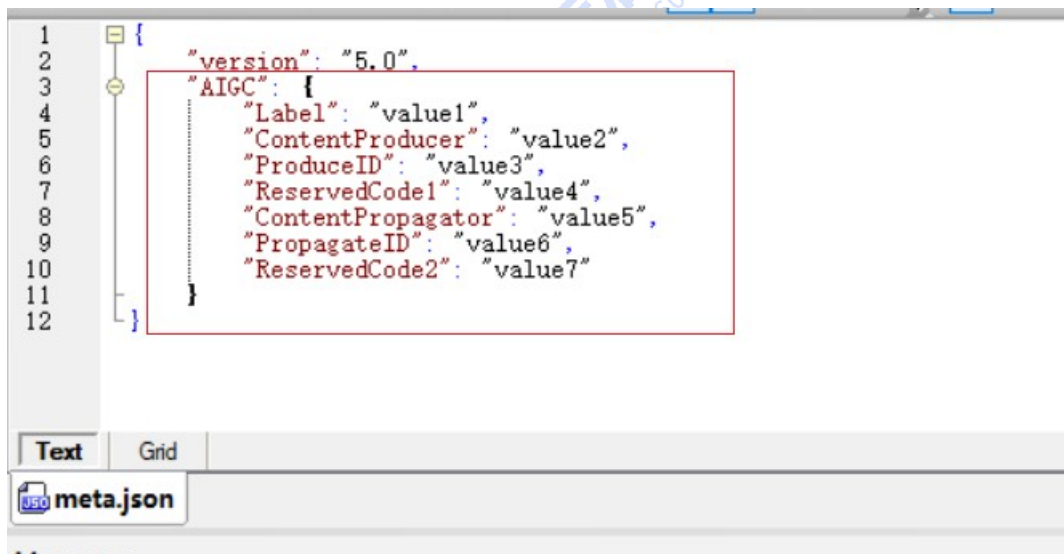


图 14 posm、posf 文件元数据标识结果示例



参考文献

- [1] ISO/IEC 29500—1 Information technology — Document description and processing languages — Office Open XML File Formats
- [2] ISO 32000-2:2020 Document management — Portable document formatPart 2: PDF 2.0
- [3] GB/T33190—2016 电子文件存储与交换格式 版式文档
- [4] GB/T44720—2024 电子文件存储与交换格式 流式文档
- [5] <https://poi.apache.org/>
- [6] <https://daringfireball.net/projects/markdown/>

