

## 中华人民共和国司法行政行业标准

SF/T 0075—2020

---

### 网络文学作品相似性检验技术规范

Technical specification for similarity examination of online literature works

2020 - 05 - 29 发布

2020 - 05 - 29 实施

---

中华人民共和国司法部

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 仪器设备.....	2
5 检验步骤.....	2
6 检验记录.....	4
7 鉴定意见.....	5
参考文献.....	6

## 前 言

本标准按照GB/T 1.1—2009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。

本标准由司法鉴定科学研究院提出。

本标准由司法部信息中心归口。

本标准起草单位：司法鉴定科学研究院、公安部第三研究所、中国科学院软件研究所、上海市人民检察院、盘石软件（上海）有限公司、上海弘连网络科技有限公司。

本标准主要起草人：郭弘、施少培、李岩、吴松洋、张颖、丁丽萍、高峰、卞新伟、孙维龙、奚建华、李毅、陆道宏、钱志高。

# 网络文学作品相似性检验技术规范

## 1 范围

本标准规定了网络文学作品相似性检验的仪器设备、检验步骤、检验记录和鉴定意见。  
本标准适用于网络文学作品的相似性检验。

## 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 29360 电子物证数据恢复检验规程  
GA/T 1568—2019 法庭科学 电子物证检验术语  
SF/Z JD0300001 声像资料鉴定通用规范  
SF/Z JD0300002—2015 音像制品同源性鉴定技术规范  
SF/Z JD0300002—2018 数字声像资料提取与固定技术规范  
SF/Z JD0301001 录音资料鉴定规范  
SF/Z JD0301003—2015 录音资料处理技术规范  
SF/Z JD0302002—2015 图像资料处理技术规范  
SF/Z JD0302003—2018 数字图像修复技术规范  
SF/Z JD0304001 录像资料鉴定规范  
SF/Z JD0401002—2015 手机电子数据提取操作规范  
SF/Z JD0402004—2018 电子文档真实性鉴定技术规范

## 3 术语和定义

GA/T 1568—2019、SF/Z JD0300001、SF/Z JD0300002—2015、SF/Z JD0300002—2018、SF/Z JD0301003—2015、SF/Z JD0302002—2015界定的以及下列术语和定义适用于本文件。

### 3.1

**网络文学作品** online literature works

以互联网为展示平台和传播媒介,以文本、图片和音视频等手段来表现的文学作品。

### 3.2

**检材** material for examination

需要进行检验的网络文学作品。

### 3.3

**样本** material for comparison

供比对的网络文学作品。

### 3.4

#### 数字水印 digital watermark

一种运用计算机算法嵌入载体文件的保护信息。

[GB/T 36303—2018, 定义3.1]

## 4 仪器设备

### 4.1 硬件

检验所用仪器设备硬件包括但不限于：

- a) 电子数据检验工作站；
- b) 声像资料检验工作站；
- c) 电子数据存储介质复制设备；
- d) 电子数据存储介质写保护设备；
- e) 数码照相机；
- f) 数码摄像机；
- g) 数码录音设备。

### 4.2 软件

检验所用仪器设备软件包括但不限于：

- a) 完整性校验值计算软件；
- b) 文件比较软件；
- c) 音视频播放软件；
- d) 截屏软件；
- e) 屏幕录像软件；
- f) 数据恢复软件；
- g) 数据分析软件；
- h) OCR（光学字符识别）文字识别工具；
- i) 语音识别工具；
- j) 网络数据包监测分析工具。

## 5 检验步骤

### 5.1 记录检材/样本的情况

5.1.1 检材/样本保存在存储介质、手机、平板电脑和电子阅览器等电子设备时，应采取以下措施：

- a) 如检材/样本的载体处于封存状态，对送检状态及拆封过程进行拍照或录像；
- b) 对检材/样本的载体进行唯一性编号；
- c) 对检材/样本的载体进行拍照或录像，记录其特征。

5.1.2 检材/样本保存在网络（包括但不限于论坛、网站、在线阅读平台、网盘和移动 APP 等）上时，宜记录检材/样本的存储位置、远程访问的账号及口令等信息。

## 5.2 数据提取与固定

### 5.2.1 载体为电子设备时，采取以下方式进行数据提取与固定：

- a) 对于具备完整性备份条件的电子设备，宜使用电子数据存储介质复制设备制作电子数据副本，并进行完整性校验。之后使用电子数据副本进行数据提取，并计算提取文件的完整性校验值；
- b) 对于具备写保护条件的电子设备，宜使用电子数据存储介质写保护设备进行数据提取，并计算提取文件的完整性校验值；
- c) 对于手机等移动终端设备，应按 SF/Z JD0401002—2015 要求进行数据提取，并计算提取文件的完整性校验值；
- d) 对于载体为只能屏幕显示或只能播放、无法提取数据的电子设备，宜使用数码照相机、数码摄像机拍摄屏幕显示的内容或使用数码录音设备录制播放的内容，将拍摄或录制获得的图像文件或音视频文件导出，并计算其完整性校验值。

### 5.2.2 载体为网络时，采取以下方式进行数据提取与固定：

- a) 对于可采用远程访问方式提取电子数据的，应启动数码摄像机或屏幕录像软件，记录提取过程，并对提取文件进行完整性校验；
- b) 对于无法采用远程访问方式提取电子数据的，可采取截屏、拍照、录音或录像等方式固定相关内容。

## 5.3 数据恢复与数据处理

数据恢复与处理要求包括：

- a) 对于被删除的检材/样本，应按 GB/T 29360 要求进行数据恢复；
- b) 对于需要进行图像增强等处理的检材/样本，应按 SF/Z JD0302002—2015 要求进行图像处理；
- c) 对于需要进行图像修复的检材/样本，应按 SF/Z JD0302003—2018 要求进行图像修复；
- d) 对于需要进行音频增强等处理的检材/样本，应按 SF/Z JD0301003—2015 要求进行音频处理。

## 5.4 阅读和下载来源分析

阅读和下载来源分析要求包括：

- a) 对于保存在电子设备上的检材/样本，可使用数据分析软件分析检材/样本的来源；
- b) 对于保存在网络上的检材/样本，可采取截屏、拍照或录像等方式记录检材/样本的来源；适用时，通过网络数据包监测，分析检材/样本的阅读或下载的来源。

## 5.5 数据一致性比对

应分别计算检材与样本文件的完整性校验值，若检材与样本文件的完整性校验值相同，则两者的数据相同；若检材与样本文件的完整性校验值不同，则根据检材与样本的实际情况，进行进一步检验。

## 5.6 文件信息的检验

检材/样本为电子文件的，可按 SF/Z JD0402004—2018 要求分别对检材与样本的文件属性和元数据等进行检验。检材/样本含数字水印的，应记录相关信息。

## 5.7 同源性比对

检材/样本为图像或音视频的，可按 SF/Z JD0300002—2015 要求进行同源性鉴定。

## 5.8 内容相似性比对

应根据检材与样本的文件类型，选择适当的方式进行内容相似性比对。内容相似性比对的方式包括但不限于：

- a) 文档比对：对检材与样本的文档目录结构、文本内容进行比对；
- b) 图像比对：
  - 1) 对检材与样本中图像内容进行比对，可按 SF/Z JD0304001 要求进行；
  - 2) 对图像中包含文字的，可使用 OCR 文字识别工具进行文字识别，经验证后，按 5.8 a) 进行比对。
- c) 音频比对：对检材与样本中的需要鉴定的音频内容进行比对，可按 SF/Z JD0301001 要求进行。也可将音频内容转换为文字后，按 5.8 a) 进行比对；
- d) 视频比对：对检材与样本中需要鉴定的视频内容进行比对，可按 5.8 b) 和 5.8 c) 进行比对。

## 5.9 结果保存

将提取或固定的数据采用封盘刻录方式刻录在空白光盘上或者保存在专用存储介质中，并计算其完整性校验值。

## 6 检验记录

### 6.1 基本要求

与检验有关的情况应及时、客观和全面地记录，保证检验过程和结果的可追溯性。

### 6.2 载体为电子设备

对于检材/样本的载体为电子设备的，应记录：

- a) 电子设备的类别；
- b) 电子设备的品牌、型号（如适用）；
- c) 电子设备的出厂时唯一性编号（如适用）；
- d) 电子设备的性状；
- e) 电子设备的照片。

### 6.3 载体为网络

对于检材/样本的载体为网络的，宜记录：

- a) 阅读/下载的地址信息；
- b) 阅读/下载检材的应用程序信息（如适用）；
- c) 用户信息。

### 6.4 提取固定过程

对于检材/样本的提取固定过程，应记录：

- a) 鉴定过程中所使用的仪器设备信息；
- b) 提取文件的元数据信息（如适用）；
- c) 提取文件的数字水印信息（如适用）；
- d) 提取文件的完整性校验值（如适用）；
- e) 远程提取的开始和结束时间（如适用）；
- f) 固定过程录像文件的文件名和完整性校验值（如适用）。



## 7 鉴定意见

### 7.1 数据相同

鉴定意见表述为：检材与样本的数据相同。

### 7.2 具有同源性

鉴定意见表述为：检材与样本具有同源性。

### 7.3 内容具有相似性

鉴定意见表述为：检材与样本中需要鉴定的内容具有内容相似性。并符合以下要求：

- a) 应列出检材与样本的相似比例，并宜对存在相同或相似的部分进行说明；
- b) 检材与样本包含多种文件类型，应分别列出各类型文件的相似比例。

注：检材与样本的相似比例包含检材与样本的相同部分占检材的比例、检材与样本的相同部分占样本的比例。

### 7.4 内容不具有相似性

鉴定意见表述为：检材与样本中需要鉴定的内容不具有相似性。

### 7.5 无法判断内容是否具有相似性

鉴定意见表述为：无法判断检材与样本中需要鉴定的内容是否具有相似性。

### 参 考 文 献

- [1] GB/T 29361—2012 电子物证文件一致性检验规程
  - [2] GB/T 36303—2018 印刷数字水印
  - [3] GA/T 754—2008 电子数据存储介质复制工具要求及检测方法
  - [4] GA/T 755—2008 电子数据存储介质写保护设备要求及检测方法
  - [5] GA/T 756—2008 数字化设备证据数据发现提取固定方法
  - [6] GA/T 976—2012 电子数据法庭科学鉴定通用方法
  - [7] SF/Z JD0400001—2014 电子数据司法鉴定通用实施规范
  - [8] SF/Z JD0403001—2014 软件相似性检验实施规范
-