XINCHECK 文本查重

产品官网：https://www.xincheck.com

联系方式：18701138792/18810760681

北京芯锋科技有限公司
Beijing XinFeng Technology Co., Ltd.

目录

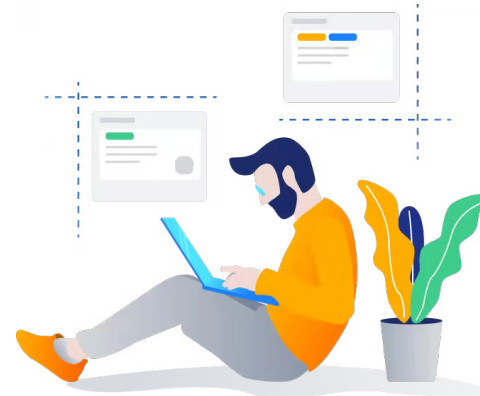
- 产品简介
- 使用教程
  - 查重 SDK 快速使用教程
  - 查重 SDK 进阶使用教程
  - 文档查重软件快速使用教程
  - 更新日志
- 采购报价
  - 查重 SDK
  - 查重软件
- Q & A

©2021 北京芯锋科技有限公司

XINCHECK 文本查重基于自研的、业界认可的查重算法及查重引擎，提供易使用、易拓展、高性能的离线查重 SDK 和查重软件，可用于论文查重、标书查重/辅助串标检测、项目申报书查重、文档查重、文本去重、作业查重等多个场景。

产品核心竞争力

1. 完全离线使用
2. 支持基于自建库查重、支持横向查重
3. 设置排除白名单，排除不需要查重的文本内容
4. 切换查重算法（文本指纹算法/分句语义算法)



©2021 北京芯锋科技有限公司

查重 SDK 主要功能

1. 完全离线使用
2. 基于自由数据构建比对库、支持横向查重
3. 线性的算法时间复杂度，任务自动并发处理，充分利用 CPU 性能
4. 生成详细的、多种类型的查重报告，且查重报告样式可以通过接口进行部分自定义
5. 支持白名单排除允许重复的内容
6. 丰富的接口，完善的开发文档，开发者可以获取查重报告中的全部数据
7. 20 余项可定制化配置项

©2021 北京芯锋科技有限公司

查重软件主要功能

1. 完全离线使用
2. 基于自由数据构建比对库、支持横向查重
3. 线性的算法时间复杂度，任务自动并发处理，充分利用 CPU 性能
4. 支持白名单排除允许重复的内容
5. 支持添加重点关注关键词
6. 支持切换查重算法（文本指纹算法/分句语义算法）
7. 支持对文档属性和图片查重并生成图片重复报告等详细的、多种类型的查重报告

©2021 北京芯锋科技有限公司

查重 SDK 快速使用教程

引用查重 SDK

通过 maven 将本 SDK 0.5.5 版本引入到项目中

```
<repository>
    <id>XINCHECK</id>
    <name>XINCHECK Public Repository</name>
    <url>https://maven.xincheck.com/repository/maven-releases/</url>
</repository>
```

然后在 <dependencies> 中添加依赖

```
<dependency>
    <groupId>com.xincheck</groupId>
    <artifactId>duplicate-check</artifactId>
    <version>0.5.5</version>
</dependency>
```

除 maven 外同样支持 Gradle、Ivy 等，修改对应引入语法即可。

©2021 北京芯锋科技有限公司

授权 SDK

我们为非商业用户提供免费的评估许可证[1]，您可先通过以下方法获取并打印服务器或 PC 机的机器指纹

```
System.out.println(CheckManager.INSTANCE.getMachineCode());
```

将机器指纹输入到下面文本框中获取许可证

机器指纹:

获取评估授权许可证

授权许可证是一串较长的包含字母数字和符号的字符串。获取后通过以下方法对 SDK 授权

```
CheckManager.INSTANCE.setRegCode("授权许可证");
```

[1] 授权时常为 3 天，可无限次续期，每次续期时长 3 天。用于教育、科研或其它合理非商业用途可申请长有效期)，详情可访问查重 SDK 采购报价表

©2021 北京芯锋科技有限公司

使用简易启动器开始查重

为了便于部分简单应用场景下的开发，SDK 0.5.0 以上版本内置了简易启动器 EasyStarter，通过如下所示的一行代码即可完成 SDK 调用。参数介绍如下

- 参数 1：**待查文件**所在的文件夹路径（如果待查文件只有一个，可以传文件路径）；
- 参数 2：**比对库文件**所在的文件夹路径（如果比对库中只有一个文件，可以传文件路径）；
- 参数 3：保存**查重报告**的文件夹路径。如果不需要导出查重报告可以传空字符串；
- 参数 4：**白名单**文本。对于标书查重等场景，有一些文本是允许重复的，这些文本可以通过该参数传入。该参数可选，如不需要可以不传或传 null。

```
List<Reporter> reporters = EasyStarter.check(new File("参数1"), new File("参数2"), "参数3", "参数4");
```

横向查重应用场景下参数 1 和参数 2 可以相同，查重时如果待查文件和比对库中的某一个文件完全相同会自动跳过比对，不会出现重复率 100

©2021 北京芯锋科技有限公司

查重 SDK 进阶使用教程

以下内容对开发过程中较为关键的部分做出介绍，详细的接口和参数说明请查阅完整开发者接入文档，完整实例代码参见 GitHub 项目中的 Sample 部分。

1. 使用原生方式同步启动任务
2. 通过 CheckState 观察者异步处理查重任务
3. 实例化文本对象（Paper）
4. 为本文对象补充额外信息
5. 为 Paper 设置 Payload

©2021 北京芯锋科技有限公司

使用原生方式同步启动任务

1 加载比对库[2]

将所有要作为比对库的文件放到一个文件夹中（支持 doc、docx、xls、xlsx、pdf、rtf、txt 格式），实例化 PaperLibrary 时将文件夹路径传入构造方法。

```
LocalPaperLibrary paperLibrary = LocalPaperLibrary.load(new File("对比库文件夹路径"));
```

将待查重的文件加载为 Paper。Paper 支持多种加载方式，可以通过文件、字符串、输入流加载，也支持批量加载。

2 加载待查重的文件

```
Paper paper = Paper.load(new File("文件路径")); //加载单个文件
```

[2] 加载比对库会花费一定时间，通常只需要在服务启动的时候加载一次，后续的所有查重操作都不需要再重新加载比对库。

©2021 北京芯锋科技有限公司

使用原生方式同步启动任务

3 启动查重任务

```
//构建并启动任务
CheckTask checkTask = CheckManager.INSTANCE
        .getCheckTaskBuilder() //获取查重任务构建器
        .addLibrary(paperLibrary) //添加比对库。可以添加多个
        .addCheckPaper(paper) //添加待查文本。可以添加多个
        .build(); //构建任务，返回checkTask对象
checkTask.start(); //启动任务线程
checkTask.join(); //等待查重结束（阻塞）
```

该方式为同步启动查重任务，并通过 join 方法等待查重任务结束。SDK 支持异步调用。

©2021 北京芯锋科技有限公司

4 保存查重报告

查重任务结束后可以通过以下方式将查重报告保存，SDK 可以生成三种查重报告。

```
checkTask.getReporters().get(0).saveAsFile("C:\\Report\\report1.html",
    ReportType.TEXT_WITH_CITATION); //保存全文标红查重报告
checkTask.getReporters().get(0).saveAsFile("C:\\Report\\report2.html",
    ReportType.TEXT_WITH_ORIGINAL); //保存原文对照查重报告
checkTask.getReporters().get(0).saveAsFile("C:\\Report\\report3.html",
    ReportType.SAMPLE); //保存简洁查重报告
```

可不将结果保存为文件，直接通过接口获取到查重结果。如:。

```
String reportId = reporter.getReportId(); //获取查重报告id
String copyRate = reporter.getCopyRate(); //获取总重复率
String copyWords = reporter.getCopyWords(); //获取重复字数
```

查重报告的样式可以通过接口进行一定程度的自定义，同时，查重报告的所有内容都可以在代码中以接口的方式获取到。

©2021 北京芯锋科技有限公司

通过 CheckState 观察者异步处理查重任务

根据比对库的大小、待查文本的字数、计算机 CPU 繁忙程度等因素的不同，需要数秒至数分钟才能完成一次查重任务。通过实现 CheckState 接口，可以实现查重任务的异步处理。

CheckState 接口中包含 start、finish、fail 三个方法，分别对应查重任务提交后的启动、完成和失败。使用 setCheckState 方法将 CheckState 注册到查重任务中。这样在查重任务的各状态就会回调 CheckState 的不同方法。

```
//构建并启动任务
CheckManager.INSTANCE
        .getCheckTaskBuilder() //获取构建者
        .setUid("1") //设置任务id。如不设置会随机生成uuid
        .addCheckState(new CheckStateImp(), "test") //设置回调处理和自定义信息。如不设置将无法收到回调
        .addLibrary(paperLibrary) //设置比对库
        .addCheckPaper(toCheckPaper) //设置待查Paper
        .build() //构建任务
        .submit(); //启动任务。submit：将任务提交到线程池中。start：直接启动任务
```


©2021 北京芯锋科技有限公司

实例化文本对象（Paper）

❶ 通过 File 对象实例化

```
Paper paper = Paper.load(new File("文件路径"));
```

❷ 通过 File 对象实例化，并指定文件类型

```
Paper paper = Paper.load(new File("文件路径"), FileType.TXT);
```

❸ 通过文本实例化

```
Paper paper = Paper.load("文本内容");
```


©2021 北京芯锋科技有限公司

为文本对象补充额外信息

对于每一个 Paper 对象，包含 id、标题、作者、其它信息四项可选信息，这四项信息的设置不会影响查重结果。id 可由开发者按需标记、使用，用以唯一标识一个 Paper 对象，不会在查重报告中展示；其余三项项信息设置后会在查重报告中进行展示。如果不设置这些信息，将默认使用文件名作为标题，其它信息为空。有以下几种方式可以对这四项信息进行设置。

1. 通过 set 方法直接设置

```
paper.setId("001").setTitle("标题").setAuthor("作者").setInfo("其它需要展示的信息");
```

2. 通过格式化文件名设置通过 File 对象加载 Paper 对象时，如果文件名中包含分隔符“@”且数量符合以下任一种规则，将自动从文件名中读取标题、作者、来源和年份信息[3]。

[3] “@”只是 SDK 的默认分割符，分隔符可由开发者自行修改，修改方式参见“详细文档-接口文档-高级配置项”部分。

©2021 北京芯锋科技有限公司

为文本对象补充额外信息

分割规则

1） 如果文件名中包含 1 个分隔符，SDK 会将文件名按分隔符分割为数组后依次读取为 id、标题，如文件名为“001@ 标题.docx”；

2） 如果文件名中包含 2 个分隔符，SDK 会将文件名按分隔符分割为数组后依次读取为 id、标题、作者，如文件名为“001@ 标题 @ 作者.docx”；

3） 如果文件名中包含 3 个分隔符，SDK 会将文件名按分隔符分割为数组后依次读取为 id、标题、作者、其它信息，如文件名为“001@ 标题 @ 作者 @ 其它需要展示的信息.docx”；

4） 如果文件名中不包含分隔符，SDK 会将文件名读取为标题，如文件名为“标题.docx”。

©2021 北京芯锋科技有限公司

为 Paper 设置 Payload

Payload 可用于传递上下文信息，也可以用于为 Paper 补充不希望在查重报告中展示的信息。对于每一个 Paper 对象，可以设置一个 Payload，Payload 不会展示到查重报告中，但可以用来存储额外信息或上下文信息。Payload 可以是任意对象，但该对象必须实现了 Serializable 接口。设置 payload 后，可以通过 get 方法获取。

```
paper.setPayload("001");
Object payload = paper.getPayload();
```

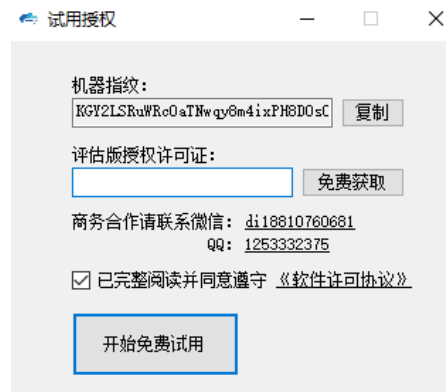

©2021 北京芯锋科技有限公司

文档查重软件快速使用教程

1. **下载**、安装[a]
2. 获取并使用免费授权许可证启动应用程序，勾选“已完整阅读并同意遵守《软件许可协议》”。点击“免费获取”自动获取授权，然后点击“开始免费试用”即可进入主程序[b]。

[a] 目前只支持 win7 及以上 64 位操作系统，不支持苹果系统

[b] 免费版支持对 8000 字以下的文档进行查重，如果超过 8000 字会被自动截断



©2021 北京芯锋科技有限公司

文档查重软件快速使用教程

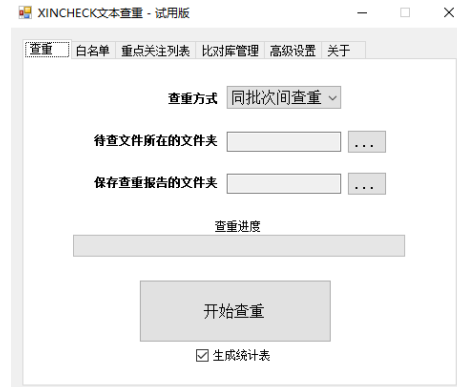
芯锋科技 XinFeng Tech.

❸ 开始查重

选择查重方式（支持同批次间查重[a]、使用比对库查重两种方式[b]）、选择待查文件所在的文件夹、保存查重报告的文件夹，点击开始查重按钮即可。

[a]同批次间查重：比较同一批次内是否存在横向抄袭的问题。

[b]使用比对库查：将文件夹中的文件和比对库中的文件比对，检查是否有与比对库中文本重复的情况。



©2021 北京芯锋科技有限公司

 文档查重软件快速使用教程

芯锋科技 XinFeng Tech.

查重完毕后会生成如下样式的查重报告：

文本复制检测报告单

№:U20200301R_15830740606238857492207067822884665633 检测时间：2020-03-01 22:47:40

检测文献：测试标题

作者：测试作者

检测结果

总文字复制比：68% 单篇最大文字复制比：38.5%

重复字数：[11410]	总字数：[16782]	单篇最大重复字数：[6468]
总段落数：[4]	前部重合字数：[537]	疑似段落最大重合字数：[7339]
疑似段落数：[4]	后部重合字数：[0]	疑似段落最小重合字数：[0]

3.2% (537) 测试标题_目录摘要等 (总3942字)

43.7% (7339) 测试标题_第1部分 (总8585字)

21.1% (3534) 测试标题_第2部分 (总4255字)

(注释：▌无问题部分 ▌文字复制比部分

1. 测试标题_目录摘要等 总字数：3942

相似文献列表 文字复制比：3.2% (537)

1	测试标题203 测试作者080 测试来源350 测试年份542	0.7% (120) 是否引证：否
2	测试标题238 测试作者157 测试来源423 测试年份229	0.6% (95) 是否引证：否
3	测试标题790 测试作者855 测试来源076 测试年份089	0.4% (59) 是否引证：否

简洁查重报告

2. 更改 (1).docx_第1部分 总字数：9165

相似文献列表 文字复制比：93.3% (8552)

1	更改.docx	93.3% (8552)
2	123(1).docx	0.2% (15)

原文内容

key words: project overview. layout. construction schedule. construction plan

目 录

第一章 综合说明

1.1 编制说明

1.我公司接受合同文件的要求，包括与项目的施工质量，施工进度，安全文明施工有关的各种控制和协调管理要求，并在领导和监督下实施各种施工计划和技术措施建筑单位。建设南通中央创新区科技创新中心一期（南通先进研究院）/ a-2项目。

2.通过仔细研究和分析施工方提供的施工图及相关技术信息，并结合现场调查，我们分析了一些可能影响施工的因素以及该项目的特点，困难和总承包。实力等各种优势，我公司有充分的信心和能力确保高质量，高速度地完成总承包管理职责，并向建设单位提交满意的答卷。

3.施工组织设计突出先进合理的施工方案，施工组织周到严谨，施工管理严谨细致，全面负责，切实做到一流，施工技术一流。科学的管理方法，一流的施工速度，一流的工程质量实现本项目的建设，达到高质量，高速度完成的目标。

4.我公司将全力做好施工现场的初步建设和生产设施的总体规划及布置的准备工作。充分利用公司的管理优势，建立健全的项目管理机构，实行严格的责任制，在建设，管理，监理单位的领导和监督下，实行项目总承包管理制度。通过对劳动力，机器设备，材料，技术，施工方法和信息的优化处理，达到了预期的施工期，质量和社会声誉的预期效果。

5.我公司严格按照施工技术规范标准和有关质量规定进行施工。特别是，它将充分执行国家颁布的强制性标准，以确保该项目一次合格并被接受并交付给业主。

6.为确保本项目的最终质量，我公司将建立严格的质量保证体系，加强施工质量验收制度，从不违反规定，从不使用不合格的材料，并真诚接受政府质量监督部门的监督。所有级别，直到项目完成验收。

7.为确保本项目安全生产，我公司在实施施工组织设计的同时编制安全生产保障计划，实行各级人员安全生产责任制，并明确项目负责人为总负责人。负责项目安全，专职安全员进行安全管理。

1.2 编制原则

1.认真执行国家工程建设的法律，法规，规章，指导方针和政策。

2.严格执行工程施工程序，坚持合理的施工程序，施工顺序和施工工艺。

3.运用现代建筑管理原则，流动施工方法和网络施工技术，组织有节奏，均衡和连续的施工。

4.优先采用先进的施工技术，科学确定施工方案；认真制定实施计划，严格控制质量，进度，成本和安全施

全文标红查重报告

4	此处有 80 字相似 务的重要技术经济文件。本施工组织设计主要包括以下内容：本工程施工组织设计书、施工平面图、施工进度计划图。用这三项内容来指导本工程施工顺利按时完成，争创"鲁班奖"或"国家优质工程"奖。关键词：工程概况，平面布	更改.docx 务的重要技术经济文件。本施工组织设计主要包括以下内容：本工程施工组织设计书、施工平面图、施工进度计划图。用这三项内容来指导本工程施工顺利按时完成，争创"鲁班奖"或"国家优质工程"奖。关键词：工程概况，平面布 更改.docx 有一个组团(1栋及以上))，争创"鲁班奖"或"国家优质工程"奖。 2 工程工期 江苏省南
5	此处有 26 字相似 "或"国家优质工程"奖。关键词：工程概况，平面布置图，施工进度计划，施工方案 construction	更改.docx "或"国家优质工程"奖。关键词：工程概况，平面布置图，施工进度计划，施工方案 construction
6	此处有 119 字相似 ，施工进度计划，施工方案 construction organization and design of no.1 basement project of nantong central innovation district. jiangsu province. pick to. c	更改.docx ，施工进度计划，施工方案 construction organization and design of no.1 basement project of nantong central innovation district. jiangsu province. pick to. c
7	此处有 255 字相似 . pick to. construction organization design is the core and soul of a project construction project construction organization. is to a project in the process of construction of all the ideas and concrete implementation plan. is a comprehensive building must document. this paper i	更改.docx . pick to. construction organization design is the core and soul of a project construction project construction organization. is to a project in the process of construction of all the ideas and concrete implementation plan. is a comprehensive building must document. this paper i
	此处有 885 字相似 t document. this paper is an engineering design	更改.docx t document. this paper is an engineering design

原文对照查重报告

©2021 北京芯锋科技有限公司

文档查重软件快速使用教程

芯锋科技 XinFeng Tech.

④ 排除部分文本

在进行标书查重或论文查重时，对于技术规格说明书、招标文件、专有名词列表中的内容，是允许重复的，XINCHECK 文档查重软件提供了白名单排除功能，将允许重复的文件或文本添加进去，则可以在最后的查重报告中进行排除，适合标书查重。

©2021 北京芯锋科技有限公司

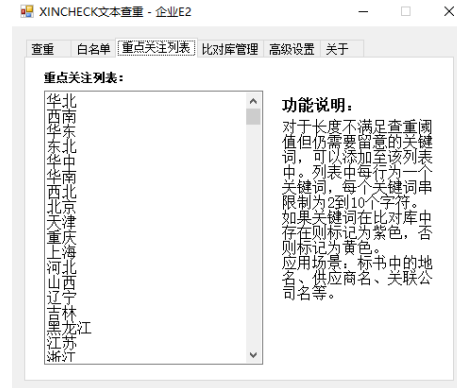
文档查重软件快速使用教程

芯锋科技 XinFeng Tech.

❺ 添加重点关注关键词

在制作标书或检查串标时，地名、供应商、关联公司等，很容易在复制时被遗漏，通过将这类关键词[a]添加至重点关注列表，在查重时将以紫色和黄色重点标注这些关键词，便于工作人员检查。

[a] 重点关注列表中的关键词长度不能超过 8 个字符



©2021 北京芯锋科技有限公司

文档查重软件快速使用教程

查重报告中紫色和黄色标注的部分即为命中重点关注列表的关键词（重复部分命中标记为紫色，非重复部分命中标记为黄色）。

2	123(1).docx	0.2%（15）

原文内容

key words: project overview, layout, construction schedule, construction plan

目 录

第一章 综合说明

1.1 编制说明

1.我公司接受合同文件的要求，包括与项目的施工质量，施工进度，安全文明施工有关的各种控制和协调管理要求，并在领导和监督下实施各种施工计划和技术措施建筑单位。建设南通中央创新区科技创新中心一期（南通先进研究院）/ a-2项目。

2.通过仔细研究和分析施工方提供的施工图及相关技术信息，并结合现场调查，我们分析了一些可能影响施工的因素以及该项目的特点，困难和总承包。实力等各种优势，我公司有充分的信心和能力确保高质量，高速度地完成总承包管理职责，并向建设单位提交满意的答卷。

3.施工组织设计突出先进合理的施工方案，施工组织周到严谨，施工管理严谨细致，全面负责，切实做到一流，施工技术一流。科学的管理方法，一流的施工速度，一流的工程质量实现本项目的建设，达到高质量，高速度完成的目标。

©2021 北京芯锋科技有限公司

文档查重软件快速使用教程

芯锋科技 XinFeng Tech.

❻ 其他功能

支持自建比对库、切换查重算法（文本指纹算法/分句语义算法）、选择查重严格性、对文档属性和图片查重并生成图片重复报告等功能。

©2021 北京芯锋科技有限公司

查重 SDK 更新日志

- 2021.11.13 发布 0.5.5 beta 版

  增加了实验特性：图片查重并生成图片重复报告。目前仅面向付费用户进行公测，需要通过高级配置项打开该功能。【评估版可以进行该项配置但功能不生效】

- 2021.11.25 发布 0.5.5

  修复若干 bug；支持基于语义的查重（实验性）。

- 2021.12.12 发布 0.5.5 patch1

  修复超过字符限制、实际不相同、前 n 字相同的文件被判定为相同文件的问题。

©2021 北京芯锋科技有限公司

查重 SDK 更新日志

- 2021.11.13 发布 0.5.5 beta 版

  增加了实验特性：图片查重并生成图片重复报告。目前仅面向付费用户进行公测，需要通过高级配置项打开该功能。【评估版可以进行该项配置但功能不生效】

- 2021.11.25 发布 0.5.5

  修复若干 bug；支持基于语义的查重（实验性）。

- 2021.12.12 发布 0.5.5 patch1

  修复超过字符限制、实际不相同、前 n 字相同的文件被判定为相同文件的问题。

©2021 北京芯锋科技有限公司

查重 SDK 更新日志

- 2021.11.13 发布 0.5.5 beta 版

  增加了实验特性：图片查重并生成图片重复报告。目前仅面向付费用户进行公测，需要通过高级配置项打开该功能。【评估版可以进行该项配置但功能不生效】

- 2021.11.25 发布 0.5.5

  修复若干 bug；支持基于语义的查重（实验性）。

- 2021.12.12 发布 0.5.5 patch1

  修复超过字符限制、实际不相同、前 n 字相同的文件被判定为相同文件的问题。

©2021 北京芯锋科技有限公司

查重软件更新日志

- 2021.10.22 发布 0.9.31
  修复一个文件属性查重的 bug。
- 2021.11.15 发布 0.9.32
  支持切换文本块指纹、分句语义两种查重算法；支持对图片查重并生成图片重复报告（实验性）。
- 2021.12.11 发布 0.9.33
  修复超过字符限制、实际不相同、前 n 字相同的文件被判定为相同文件的问题；修复横向查重多个相同文件只生成一份查重报告的问题。

©2021 北京芯锋科技有限公司

查重软件更新日志

- 2021.10.22 发布 0.9.31
  修复一个文件属性查重的 bug。
- 2021.11.15 发布 0.9.32
  支持切换文本块指纹、分句语义两种查重算法；支持对图片查重并生成图片重复报告（实验性）。
- 2021.12.11 发布 0.9.33
  修复超过字符限制、实际不相同、前 n 字相同的文件被判定为相同文件的问题；修复横向查重多个相同文件只生成一份查重报告的问题。

©2021 北京芯锋科技有限公司

查重软件更新日志

- 2021.10.22 发布 0.9.31
  修复一个文件属性查重的 bug。
- 2021.11.15 发布 0.9.32
  支持切换文本块指纹、分句语义两种查重算法；支持对图片查重并生成图片重复报告（实验性）。
- 2021.12.11 发布 0.9.33
  修复超过字符限制、实际不相同、前 n 字相同的文件被判定为相同文件的问题；修复横向查重多个相同文件只生成一份查重报告的问题。

©2021 北京芯锋科技有限公司

查重 SDK

许可证版本	企业 E1	企业 E2	企业 E3	企业 E4
授权时长	当前版本永久授权	当前版本永久授权	当前版本永久授权	当前版本永久授权
发布地点数	1 个 支持离线部署	同一地级市内 3 个 支持离线部署	同一地级市内不限 支持离线部署	不限地级市、发布地点 支持离线部署
测试服务器	数量不限	数量不限	数量不限	数量不限
比对库限制	总字数 5 亿以内	总字数 30 亿以内	均不限	均不限
待查文本限制	均不限	均不限	均不限	均不限
查重报告样式	支持自定义	支持自定义	支持自定义	支持自定义
高级配置项及特性	支持全部	支持全部	支持全部	支持全部
水印及广告	无	无	无	无
软件更新	1 年	1 年	1 年	1 年
技术支持	180 天邮件支持 30 天在线技术支持	1 年邮件支持 45 天在线技术支持	1 年邮件支持 45 天在线技术支持	1 年邮件支持 90 天在线技术支持
价格	￥12800	￥29800	￥49800	￥198800

* **续费优惠策略**：软件更新服务到期前续费享 5 折优惠；软件更新服务到期后续费享 6 折优惠。

©2021 北京芯锋科技有限公司

查重软件

许可版本	个人 P1	企业 E1	企业 E2	企业 E3
授权时长	当前版本永久授权	当前版本永久授权	当前版本永久授权	当前版本永久授权
发布地点数	1 个 支持内网使用	同一地级市内 3 个 支持内网使用	同一地级市内 30 个 支持内网使用	不限地级市 发布地点数 3000 个 支持内网使用
比对库限制	总字数 400 万以内 篇数上限 30	总字数 5 亿以内 篇数上限 1000	均不限	均不限
待查文本限制	单篇 40 万字以内 单次查重篇数 10 以内、400 万字内	单篇不限 单次查重篇数 20 以内	均不限	均不限
水印及广告	无	无	无	无
软件更新	1 个月	1 年	1 年	1 年
技术支持	30 天邮件支持	180 天邮件支持 1 次使用培训 30 天在线技术支持	1 年邮件支持 1 次使用培训 30 天在线技术支持	1 年邮件支持 1 次使用培训 90 天在线技术支持
价格	￥1980	￥12800	￥29800	￥198800

* **当前版本永久授权**：技术支持周期内包含 bug 修复更新，但不包含大版本更新及周期结束后的任何更新。

©2021 北京芯锋科技有限公司

Q & A Thank you

©2021 北京芯锋科技有限公司

芯锋科技 XinFeng Tech.

1 Question

- 查重性能如何?
  - 内存占用：SDK 启动后会占用约 50MB 常驻内存。比对库中每一千万字次需要约 150MB 常驻内存。此外，在加载比对库时解析 word 和 pdf 文档时所需的动态内存约 1 2GB，动态内存在 GC 后会被回收。
  - CPU 占用：在没有任何任务运行时，SDK 不会占用 CPU。在进行比对库导入时，如果开启了多线程导入通常会占用约 50% 的 CPU，如果未开启多线程导入通常会占用一个 CPU 核心。在进行查重任务时，每个待查文本的文本段将占用一个 CPU 核心，SDK 默认将每 1 万字划分为一个文本段。如果您需要预留部分服务器的 CPU 资源用于其它服务，可以对 SDK 进行并发配置，限制 SDK 的 CPU 占用。
  - 查重速度：待查文本每 1 万字、比对库每 1 千万字次的情况下秒级出结果。时间随数据量线性增加。

©2021 北京芯锋科技有限公司

芯锋科技
XinFeng Tech.

2 Question

- SDK 的发布地点是什么意思?

  终端的定义为:“具备存储功能且可以进行逻辑计算的实体或虚拟设备，包括但不限于服务器、台式电脑、手提电脑、手持移动终端、电视机、机顶盒、单片机、开发板等，以及使用虚拟化技术生成的上述实例（如虚拟机等)。”发布地点的定义为:“嵌入许可软件的应用程序用于商业用途时的分发地点。如果是桌面应用程序 (C/S)，每个终端均被视为一个发布地点。若是服务器应用程序，分发到内网或私有云平台，且许可软件提供的能力通过公网不可以直接或间接访问，则每个终端均被视为一个发布地点；分发到公有云平台（如：阿里云、腾讯云、华为云等)，且许可软件提供的能力通过公网可以直接或间接访问，则被视为无限发布地点；分发到有公云平台且许可软件提供的能力无法通过公网直接或间接访问，则每个终端被视为一个发布地点。

©2021 北京芯锋科技有限公司

芯锋科技
XinFeng Tech.

3 Question

- 桌面端的发布地点数是什么意思?

  可以简单理解为装机数量（即安装到几台电脑上）。详细定义如下：

  终端的定义为：“具备存储功能且可以进行逻辑计算的实体或虚拟设备，包括但不限于服务器、台式电脑、手提电脑、手持移动终端、电视机、机顶盒、单片机、开发板等，以及使用虚拟化技术生成的上述实例（如虚拟机等）。”

  发布地点的定义为：“每一个安装 XINCHECK 文档查重桌面端软件的终端即为一个发布地点。如果终端的主要硬件（主板和网卡）发生变化，则视为产生了一个新的发布地点。”

©2021 北京芯锋科技有限公司

芯锋科技 XinFeng Tech.

4 Question

- 在哪些领域有过合作案例?
  有过电子招标系统、科研管理平台、论文查重系统、舆情系统等合作案例。

©2021 北京芯锋科技有限公司