**上传**

POST /upload

下载

GET /download

参数filename: String

**隐私动态度量工具集**

**要求：**具备隐私的分量量化、分量组合与重度量、动态调整等功能，支持数据类型≥10类

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 数据类型 | 语言 | 说明 | API | 参数 | 返回 | 样例 |
| 1 | 数值表 | Python | LLM | POST /measure/number | num\_to\_measure: Array of Number 待测数值数组  num\_private: Number 敏感数字 | result: Bool. 是否存在隐私信息（敏感数字是否在待测数组中） | [100, 1000]  100 |
| 2 | 图片 | Python | LLM | POST /measure/image | filename: String 待测文件名 | result: Float. 存在隐私信息的概率 | face.jpg |
| 3 | 视频 | Python | LLM | POST /measure/video | filename: String 待测文件名 | result: Float. 存在隐私信息的概率 | test.avi |
| 4 | 音频文件 | Python | LLM | POST /measure/audio | filename: String 待测文件名  keyword: String 敏感关键词（可选，默认为空字符串） | result: Float. 存在隐私信息的概率，或keyword不为空时敏感关键词出现的概率 | test.flac  空，或"America" |
| 5 | 实时语音 | Python | LLM | POST /measure/audio\_stream | filename: String 待测文件名  keyword: String 敏感关键词（可选，默认为空字符串） | result: Float. 存在隐私信息的概率，或keyword不为空时敏感关键词出现的概率 |  |
| 6 | pdf | Python | LLM | POST /measure/pdf | filename: String 待测文件名 | result: Float. 存在隐私信息的概率 | ori\_pdf.pdf |
| 7 | csv | Python | LLM | POST /measure/csv | filename: String 待测文件名 | result: Float. 存在隐私信息的概率 | ori\_csv.csv |
| 8 | 文字 | Python | LLM | POST /measure/text | text: String 待测文本  keyword: String 敏感关键词（可选，默认为空字符串） | result: Float. 存在隐私信息的概率，或keyword不为空时敏感关键词出现的概率 | "I went to the park and saw a dog."  "go" |
| 9 | 位置 | Python | LLM | POST /measure/location | location: Array of Float 待测坐标（经纬度二元数组）  zone\_coords: Array of Array of Float 敏感区域圆心的坐标构成的数组，至少有一个坐标  dist\_thresh: Float 敏感区域的半径（米），（可选）默认1000 | result: Bool. 是否存在隐私信息（待测坐标是否在敏感区域中） | location: (116.435842, 39.941626)  zone\_coords: [(116.435842, 39.941626), (116.353714, 39.939588), (116.435806, 39.908501), (116.356866, 39.907242)] |
| 10 | 轨迹 | Python | LLM | POST /measure/trace | trace: Array of Array of Float 待测轨迹（坐标数组）  zone\_coords: Array of Array of Float 敏感区域圆心的坐标构成的数组，至少有一个坐标  dist\_thresh: Float 敏感区域的半径（米），（可选）默认1000 | result: Bool. 是否存在隐私信息（待测轨迹的任意一点是否在敏感区域中） | trace: [(1.0, 2.0), (116.435842, 39.941626), (3.0, 4.0)]  zone\_coords: [(116.435842, 39.941626), (116.353714, 39.939588), (116.435806, 39.908501), (116.356866, 39.907242)] |
| 11 | ofd | Python | LLM | POST /measure/ofd | filename: String 待测文件名 | result: Float. 存在隐私信息的概率 | ori\_ofd.ofd |

**隐私保护效果评估工具集**

**要求：**具备单算法/多算法组合的保护效果评估、多因素平衡的保护效果评估等功能，支持数据类型≥10类

目前已跑通4类共11个算法【均为Python，命令行调用，代码共345行】：

数值表 1个：信息损失评估DP

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 数据类型 | 语言 | 说明 | API | 参数 | 返回 | 样例 |
| 1 | 数值表 | Python | 信息损失（联合熵做差） | POST /assess/table | filename1: String 脱敏前文件名(csv)  filename2: String 脱敏后文件名(csv) | result: Float 信息损失 | adult\_small.csv  adult\_k3.csv |
| 2 | 图片 | Python | SSIM | POST /assess/image | filename1: String 脱敏前文件名  filename2: String 脱敏后文件名 | result: Float SSIM相似度 | lena.jpg  lena\_protected.jpg |
| 3 | 视频 | Python | 每一帧SSIM算平均值。 | POST /assess/video | filename1: String 脱敏前文件名  filename2: String 脱敏后文件名 | result: Float SSIM相似度 | test.mp4  pixelate\_region.mp4 |
| 4 | 音频文件 | Python | 余弦相似度 | POST /assess/audio | filename1: String 脱敏前文件名  filename2: String 脱敏后文件名 | result: Float 余弦相似度 | audiotest.wav  beeped.wav |
| 5 | 实时语音 | Python | 余弦相似度。复用 | POST /assess/audio\_stream | filename1: String 脱敏前文件名  filename2: String 脱敏后文件名 | result: Float 余弦相似度 |  |
| 6 | pdf | Python | 语意相似性。神经网络。 | POST /assess/pdf | filename1: String 脱敏前文件名  filename2: String 脱敏后文件名 | result: Float 相似度 | ori\_pdf.pdf  protect\_pdf.pdf |
| 7 | csv | Python | 语意相似性。神经网络。 | POST /assess/csv | filename1: String 脱敏前文件名  filename2: String 脱敏后文件名 | result: Float 相似度 | ori\_csv.csv  protect\_csv.csv |
| 8 | 文字 | Python | 语意相似性。神经网络。 | POST /assess/text | text1: String 脱敏前字符串  text2: String 脱敏后字符串 | result: Float 相似度 | "hello my name is john"  "hello my name is allen" |
| 9 | 位置 | Python | 语意相似性 + 数值相似性。神经网络。 | POST /assess/location | loc1: Array of Float 脱敏前坐标  loc2: Array of Float 脱敏后坐标 | result: Float 相似度 | (116.435842, 39.941626)  (1.0, 2.0) |
| 10 | 轨迹 | Python | 语意相似性 + 数值相似性 + 位置关联性。神经网络。 | POST /assess/trace | trace1: Array of Array of Float 脱敏前轨迹  trace2: Array of Array of Float 脱敏后轨迹 | result: Float 相似度 | [(1.0, 2.0), (116.435842, 39.941626), (3.0, 4.0)]  [(1.0, 2.0), (116.43794199999999, 39.932226), (3.0, 4.0)] |
| 11 | ofd | Python | 语意相似性。神经网络。 | POST /assess/ofd | filename1: String 脱敏前文件名  filename2: String 脱敏后文件名 | result: Float 相似度 | ori\_ofd.ofd  protect\_ofd.ofd |

**隐私按需脱敏工具集**

**要求：**具备脱敏控制、协同脱敏、按需脱敏等功能，支持数据类型≥10类

（所有python、C代码均通过命令行调用，所有java代码可通过命令行/jar包/IM调用）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 数据类型 | 语言 | 说明 | API | 参数 | 返回 | 样例 |
| 1 | 数值表 | Python | 对csv的每个数字脱敏，方法为把数字的每一位随机替换 | POST /desensitize/number | ifilename: String 输入文件名  ofilename: String 输出文件名 | result: String脱敏后文件下载链接，如/download?filename=protect\_table.csv。后同。 | ori\_table.csv  protect\_table.csv |
| 2 | 图片 | Python | 图片指定区域高斯模糊化 | POST /desensitize/image | ifilename: String 输入文件名  ofilename: String 输出文件名  region: Array of float 脱敏区域四元组，依次为(x, y, w, h)  radius: Float 高斯核标准差。（可选）默认=5 | result: String脱敏后文件下载链接 | lena.jpg  lena\_protected.jpg  [230, 240, 115, 130] |
| 3 | 视频 | Python | 视频指定区域像素化 | POST /desensitize/video | ifilename: String 输入文件名  ofilename: String 输出文件名  region: Array of float 脱敏区域四元组，依次为(x, y, w, h)  block\_size: Int 像素块大小。（可选）默认为5 | result: String脱敏后文件下载链接 | test.mp4  pixelate\_region.mp4  [300, 100, 500, 500] |
| 4 | 音频文件 | Python | 正弦波替换音频（哔声） | POST /desensitize/audio | ifilename: String 输入文件名  iformat: String 输入文件格式(如wav, mp3)  ofilename: String 输出文件名  oformat: String 输出文件格式  start: Float脱敏起始时间（秒）  duration: Float 脱敏持续时间（秒） | result: String脱敏后文件下载链接 | audiotest.wav  beeped.wav |
| 5 | 实时语音 | Python | 正弦波替换音频（哔声）  脱敏按钮按住表示脱敏，松开表示结束。按钮按下时得到对应的脱敏起始时间，按钮松开时得到脱敏持续时间。 | POST /desensitize/audio\_stream | ifilename: String 输入文件名  iformat: String 输入文件格式（如wav）  ofilename: String 输出文件名  oformat: String 输出文件格式  start: Float脱敏起始时间（秒）  duration: Float 脱敏持续时间（秒） | result: String脱敏后文件下载链接 |  |
| 6 | pdf | Python | 页面删除 | POST /desensitize/pdf | ifilename: String 输入文件名  ofilename: String 输出文件名  pages: Array of Int 待脱敏页码。（从1开始） | result: String脱敏后文件下载链接 | ori\_pdf.pdf  protect\_pdf.pdf  [1,2] |
| 7 | csv | Python | k匿名 | POST /desensitize/csv | ifilename: String 输入文件名  ofilename: String 输出文件名  k: int  QI\_INDEX: Array of int 准标识符列号（从0开始）  SA\_INDEX: int 敏感列列号 | result: String脱敏后文件下载链接 | adult\_small.csv  adult\_k3.csv  3  [0, 1, 2, 3, 4, 5, 6, 7]  8 |
| 8 | 文字 | Python | 关键词替换为随机字符串 | POST /desensitize/text | keyword: String 待替换的关键词  text: String 可能包含关键词的整个句子 | result: String 脱敏后字符串 | hello my name is barry  barry |
| 9 | 位置 | Python | 对经纬度坐标添加差分隐私拉普拉斯噪声 | POST /desensitize/location | loc: Array of Float 坐标（经纬度数组）  eps: Float 隐私预算（可选）。默认0.9 | result: Array of Float 脱敏后坐标（经纬度数组） | [10.3, 10.3]  0.1 |
| 10 | 轨迹 | Python | 指定脱敏区域，在区域内时对经纬度坐标添加噪声。 | POST /desensitize/trace | trace: Array of Array of Float 待脱敏轨迹（坐标数组）  zone\_coords: Array of Array of Float 敏感区域圆心的坐标构成的数组，至少有一个坐标  dist\_thresh: Float 敏感区域的半径（米），（可选）默认1000  eps: Float 隐私预算（可选）。默认0.9 | result: Array of Array of Float 脱敏后轨迹（经纬度坐标数组） | trace: [(1.0, 2.0), (116.435842, 39.941626), (3.0, 4.0)]  zone\_coords: [(116.435842, 39.941626), (116.353714, 39.939588), (116.435806, 39.908501), (116.356866, 39.907242)] |
| 11 | 数值 | Python | 对每一位数字随机替换 | POST /desensitize/number | num: Float 待脱敏数值 | result: Float 脱敏后数值 | 123.456 |
| 12 | ofd | Java | 页面删除 | POST /desensitize/ofd | ifilename: String 输入文件名  ofilename: String 输出文件名  pages: Array of Int 待脱敏页码。（从1开始） | result: String脱敏后文件下载链接 | ori\_ofd.ofd  protect\_ofd.ofd  [1,2] |