

wtsensor.com 404 清单 — 服务商交接文档

数据范围: 2026-08-18 ~ 2026-08-31 日志(213,862 请求) **404 总量**: 32,414 次(占全部请求 15.2%) **去重路径**: 5,661 个 **其中站点自身问题**: 1,916 个路径 / 14,980 次(需处理) **其中扫描攻击噪音**: 3,745 个路径 / 17,434 次(可忽略)

交接方式: 完整明细见 `404_site_issues.csv` (1,916 行,Excel 可直接打开,UTF-8-SIG 编码)。本清单按优先级列出必须处理的类别,服务商按类别核对即可。

A. 必须处理 — 模板/代码 Bug(最高优先)

A1. 「ROOT_URL」图片链接未替换(约 7,000 次 404)

首页及部分产品页 HTML 中,图片模板变量 `ROOT_URL` 未被替换,直接输出字面量:

```
/ROOT_URLdata/watermark/20260205/69843f0b9f68c200_200.webp 553 次
/ROOT_URLdata/watermark/20250530/683947976540e200_200.webp 551 次
/ROOT_URLdata/watermark/20250114/6785f86000e6c200_200.webp 551 次
/ROOT_URLdata/watermark/20241219/6763d3db7d0d0200_200.webp 551 次
...(共 14 个文件,每个 528~553 次)
```

- 特征: URL 以 `/ROOT_URLdata/` 开头,正确路径应为 `/data/`。
- referer 全部为 `https://www.wtsensor.com/` (首页),说明是**首页 HTML 直接输出的坏链接**,真实用户每次打开首页都会触发。
- 另发现产品页 `/diffused-silicon-pressure-sensors/ROOT_URLdata/...` 同类问题(每个 4~6 次)。
- 处理: ① 后台模板中查找 `ROOT_URL` 未替换的图片输出语句并修正;② 临时兜底:Nginx 加 `rewrite ^/ROOT_URLdata/(.*)/data/$1 permanent;` (或内部重写)避免迁移期间持续 404。

A2. 已删除/失效产品页面仍被 sitemap.xml 引用(约 40+ 个 URL)

sitemap.xml 中引用的页面返回 404,典型:

```
/dicing-saw/Fully Automatic Dicing Saw DS616 Supplier.html 22 次
/pressure-transmitter-with-flange/PCM2051-F/R Intelligent Flush Diaphragm Monocrystalline Silicon Flush... 19 次
/dicing-saw/Precision Dicing Saw DS9100.html 18 次
/dicing-saw/Super Precision Dicing Saw DS623.html 18 次
/pressure-sensor-with-the-thread/PC12 Flush Diaphragm Pressure Sensor with thread.html 18 次
/dicing-saw/Precision Dicing Saw DS820.html / DS9120.html / DS830.html 各 16~17 次
/iot-pressure-transmitter/pcm9870 / pcm9880 / nb-wireless / pcm86-wireless... 各 16~17 次
/latest-news/...(约 8 篇新闻) 各 15~17 次
```

- 完整清单见 CSV(category=site 且 top_referer 含 `sitemap.xml` 的行)。
- 处理: 新站 sitemap 重做时剔除这些 URL;若有替代页面(同系列新型号),做 301。

A3. 站内页面引用了不存在的资源

```
/themes/simplebootx/picbanners/web27.jpg 15 次(referer: 定制开发页,页面引用的 banner 图已删)
/themes/simplebootx/wotian/css/nav.css.map 14 次(开发期 sourcemap,生产环境不应引用)
/themes/simplebootx/wotian/css/baguetteBox.css.map 14 次
```

/data/watermark/20200122/*.jpg|png 各 2 次(旧图片被删,页面仍引用)
/static/images/nopic.png 10 次(占位图路径不存在)

- **处理:** 删除页面中对已删图片的引用; .map 文件可在生产移除或返回 200 空文件;检查 nopic.png 占位图配置。

A4. 废弃路由仍被站内引用

/index.php/Search/Index/news.html 781 次(referer: www.wtsensor.com)

- 站内搜索"新闻"功能的路由已失效。其中 389 次来自单一 IP(176.164.35.23)、218 次来自 130.49.215.90(疑似扫描),但 referer 指向站内,建议确认搜索页是否还有入口。
- **处理:** 搜索功能若已下线,把入口从页面移除;路由保留 404 即可,或用 301 指向新闻列表页。

B. 建议处理 — 有真实引用的死链(迁移时做 301 或清理)

以下 URL 来自站内页面、搜索引擎或第三方平台引用,迁移后应做 301 到对应新地址(或同系列产品页),避免丢权重:

含未转义字符(空格/&)的用户输入 URL:

/pressure-transmitter/industrial-pressure&Level-transmitter.html 7 次
/about us/contact-wt/ 与 /About+us/contact-wt/ 5+5 次
/pressure-sensor/temperature and pressure integrated sensor.html 2 次
/glass to metal seal 1 次

第三方/旧平台链接(阿里系 pd/pl 开头的旧商品页,部分有外站 referer):

/pd6xxxxx.html / pl5xxxxx.html / id3761565.html / ic45822.html 等(约 20 个,各 1~4 次)

俄语子站路径被直接访问(ru.wtsensor.com 误当路径):

/ru.wtsensor.com/...(约 10 个,各 1~2 次)

其他有站内 referer 的失效页: 详见 CSV(category=site 且 count≥3 的行,约 100 个)。

提示: 迁移时建议导出 CSV 中 count≥3 的 site 类路径,与新站 URL 对照,生成 301 映射表。

C. 忽略 — 扫描/攻击噪音(约 17,434 次,占 404 的 54%)

典型探测(均为攻击者行为,返回 404 正确,无需处理):

/.env 及其变体(127+ 次)、/.git/config(90)、/admin.php(82)、/info.php(76)
/wp-login.php(69)、/graphql(65)、/chosen.php(64)、/1.php(62)、/webmail/(59)
/.aws/credentials(50)、/serviceAccountKey.json(32)、/private-key(28)

- 请求方多为 GCP 段 IP 伪造 DeepSeekBot/Googlebot UA 的扫描器。
- 唯一安全提示: /admin 后台登录页返回 200 且已被扫描器发现,建议迁移后对后台增加 IP 白名单或二次认证(与 404 无关)。

D. 附: 迁移检查项速查

#	检查项	归属
1	首页/产品页 ROOT_URLdata 坏链接修复	A1
2	sitemap.xml 去除失效 URL、补充新 URL	A2
3	页面内失效图片/占位图/banner 引用清理	A3
4	搜索新闻路由入口下线或 301	A4
5	有外链的旧 URL 生成 301 映射(CSV count≥3)	B
6	后台登录加固(白名单/二次验证)	C
7	迁移后全站 404 复测(建议新站上线后 2 周内抓一次日志对比)	—

附: 文件清单

文件	内容
404_site_issues.csv	站点自身问题 1,916 行(path / count / top_referer / top_ua / 独立IP数 / 示例query)
404_full_list.csv	全部 5,661 行(含扫描噪音,已标 category: site / probe)
wtsensor_log_report.md	上一份整体体检报告