

日志数据格式

webplog.dat (查询日志)

每 7 行一条记录。

内容	说明
2594227161.144.60.114881287	session id
Mon Oct 20 14:11:12 2003	date
61.144.60.114	client ip
计算机动画技术的发展	query word
5	rank
1	page number
Database	cached or not

webpurl.dat (点击日志)

每 7 行一条记录。

内容	说明
17323874211.68.75.61561050	session id
Sat Jul 12 20:21:19 2003	date
211.68.75.61	client ip
撒哈拉 荒漠化	query word
http://www.llas.ac.cn/cwais/dq990220.htm	hit url
264905733	document id
39	rank