

易宝网络舆情监控应用系统 v2.0

软件产品说明书

易宝电脑系统(北京)有限公司

目录

1	产品说明.....	3
2	产品架构.....	4
3	产品功能.....	5
3.1	舆情采集.....	5
3.1.1	支持网页结构自动分析.....	6
3.1.2	支持基于模板的元数据解析.....	6
3.1.3	内嵌 Javascript 脚本执行引擎.....	6
3.1.4	支持 RSS 解析.....	7
3.1.5	支持境外网站监控.....	7
3.2	舆情信息存储和智能检索.....	7
3.2.1	中文分词.....	7
3.2.2	内容相似性分析.....	8
3.2.3	基于规则自动分类.....	8
3.2.4	基于内容自动分类.....	8
3.3	舆情管理.....	8
3.3.1	舆情信息分类.....	8
3.3.2	舆情用户、组织管理.....	8
3.3.3	舆情统计、分析.....	9
3.4	舆情服务.....	10

1 产品说明

舆情是指在一定的社会空间内，围绕社会事件的发生、发展和变化，民众对社会管理者产生和持有的社会政治态度，它是较多群众关于社会中各种现象以及问题所表达的信念、态度、意见和情绪等表现的总和。随着互联网发展的日益深入，网络已成为当今社会信息传播的主要媒介之一，网络舆情迅速形成，对社会影响巨大，不仅需要各级党政干部密切关注，也值得社会各界高度重视。

为了进一步加强互联网新闻宣传和信息内容的安全管理，各级政府部门迫切希望及时了解舆论动向，引导舆论发展，制定出正确的策略方针，并能及时采取措施进行有效的干预。因此，基于互联网的舆情监控对于了解社情民意，缓解舆论压力，建设和谐社会具有重要作用和意义。

易宝网络舆情监控应用系统（eOMS）能够实时监测跟踪网络舆情，为政府机关、企事业单位提供真实、可靠、实时的舆情监测、分析和预警。eOMS 内置监测范围覆盖了国内数百个主流媒体站点和政府站点，上百家论坛、贴吧、博客，实时并行监测、汇总和分析互联网上的舆论热点。系统具有灵活的定制性和可扩展性，可根据用户需求进行量身定制。

eOMS® 将帮助您快速全面地了解整个互联网对企业、政府单位和舆论焦点报道和讨论的历史演变、第一时间抓住突发事件、全自动地统计分析和报告舆情。

eOMS® 内核采用先进的海量文本搜索、检索技术，具有每秒钟超过数百个新增舆情文档的吞吐量，可以并行快速聚合分析和统计超过超过 3 年的论坛讨论信息，可以同时索引查找超过 T 级数量的舆情信息。主要功能有：

（一）舆情采集功能：eMOS 舆情采集工具在支持传统网页采集功能的基础上，实现了对论坛、博客、新闻评论等舆情信息的精准和全面采集，为舆情分析提供了丰富的数据资源。

（二）舆情检索功能：eMOS 全文数据库检索系统提供了强大的检索功能，可以对海量舆情数据进行高效、准确、实时的搜索。

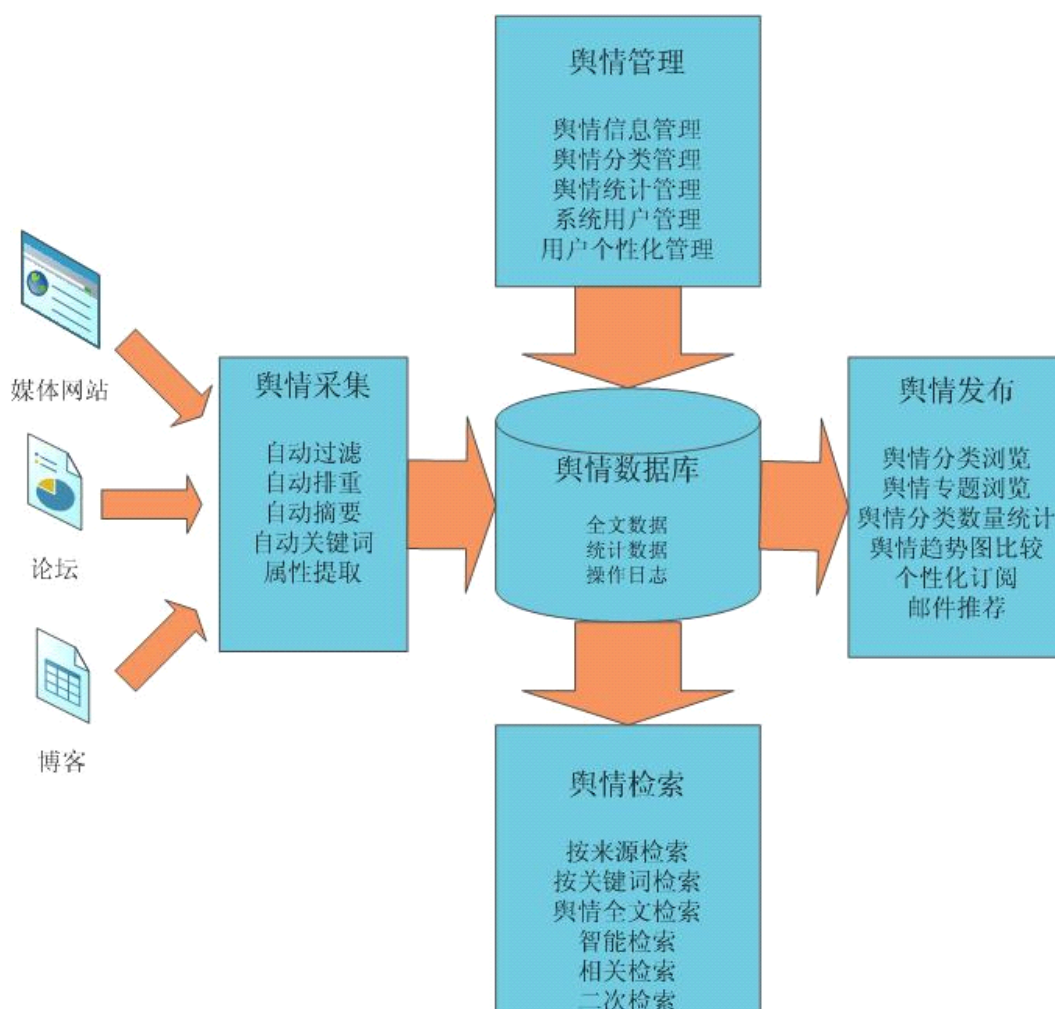
（三）舆情分析功能：eMOS 提供工具使得用户可以自行设计统计内容和输出结

果。对统计数据的排名、对时间、来源和关键词组的三维聚合分类，让您方便地从不同视角透视统计数据，直接获得不同条件下的聚合统计结果。极大程度便利了数据分析人员的操作。动态统计图、统计表、报告生成和专业的自动分析功能，也使得数据分析和专业报告撰写不再复杂。

（四）舆情展示功能：eMOS 提供灵活的个性化定制功能，用户可以根据自己需求配置信息分类，首页的显示内容、样式、摆放的位置、信息条数等，方便用户对系统的使用。还可以根据用户身份不同显示不同的分类栏目、统计内容等。

2 产品架构

eMOS 舆情监控系统由舆情采集平台、舆情发布平台、舆情管理平台和舆情检索引擎四部分组成。系统架构如下图所示：



3 产品功能

3.1 舆情采集

在网络舆情采集方面，eMOS 舆情监控系统综合采用多种先进技术打造的舆情采集工具，实现了对各种舆情信息的精准和全面采集。系统内置了主流论坛、博客、贴吧、新闻媒体、政府网站的数百个监控信息源，并对这些信息源列表定期进行更新，实现用户即买即用，节省了用户部分实施部署很大一部分成本。用户还可以根据自己的需求定制采集源。

线程	编号	URL名称	URL标题
BBS_天涯社区	0	client.zmfish.com/api.php?module=redirect&id=959&url=...	你是否真正了解韩国? - 纵论天下 - 环球论坛 - Power...
BBS_百度2	1	bbs.huangqu.com/zongluntianxia/thread-157859-5-1.html	你是否真正了解韩国? - 纵论天下 - 环球论坛 - Power...
BBS_百度3	0	client.zmfish.com/api.php?module=redirect&id=775&url=...	你是否真正了解韩国? - 纵论天下 - 环球论坛 - Power...
BBS_搜狐社区	1	bbs.huangqu.com/zongluntianxia/thread-157859-4-1.html	你是否真正了解韩国? - 纵论天下 - 环球论坛 - Power...
BBS_腾讯论坛	1	bbs.huangqu.com/zongluntianxia/thread-157859-3-1.html	你是否真正了解韩国? - 纵论天下 - 环球论坛 - Power...
BBS_海扑	4	forum.xaonline.com/viewthread.php?tid=1551909&extra=sp...	大家把西安的坏交警拿出来晒晒
BBS_中青论坛	1	bbs.huangqu.com/zongluntianxia/thread-157859-2-1.html	你是否真正了解韩国? - 纵论天下 - 环球论坛 - Power...
BBS_凤凰网论坛	1	bbs.huangqu.com/zongluntianxia/thread-157859-1-1.html	你是否真正了解韩国? - 纵论天下 - 环球论坛 - Power...
BBS_女友社区	1	bbs.huangqu.com/zongluntianxia/thread-209880-6-1.html	韩剧的精华内容是否助长了你的厌警情绪 - 纵论天下 - 环...
BBS_强国社区	1	bbs.yynet.com/forumdisplay.php?fid=1	读事读非 - 青年论坛_北青网
BBS_同济论坛	4	forum.xaonline.com/viewthread.php?tid=1553695&extra=sp...	震惊! 宝鸡公交车司机单挑王奔死!
BBS_烟台论坛	2	bbs.huanglong.com/forum-103-1.html	龙渊茶座 - 化龙卷 常州人上化龙卷 常州第一人气社...
BBS_网易论坛	1	bbs.yynet.com/viewthread.php?tid=1017863&extra=page%K3D1	晒晒人人都可能遭遇的比《蜗居》还窝心的几件事儿
BBS_新浪论坛	4	forum.xaonline.com/viewthread.php?tid=1554205	意外走失! 天鹰消防队
BBS_MSN中国社区	0	bbs.zmfish.com/thread-170351-1-1.html	[07-05] 小金鱼市修车和道具购买使用说明 - 『疑问建议...』
BBS_北方论坛	1	bbs.yynet.com/viewthread.php?tid=978745&extra=page%K3D1	「原创」“国富民强”还是“官富民穷”?
BBS_齐鲁热辣社区	1	bbs.yynet.com/viewthread.php?tid=997344&extra=page%K3D1	你赞成么? 专家呼吁一对夫妇生俩娃: 结束一子化时代
BBS_椰粉社区	3	bbs.ngzb.com.cn/viewforum.php?fid=72	南国早报论坛 - 南国早茶
BBS_浙江都市论坛	4	forum.xaonline.com/viewthread.php?tid=1551909&extra=sp...	大家把西安的坏交警拿出来晒晒 - 西安聚焦 - 古城茶秀 ...
BBS_强国社区	0	bbs.zmfish.com/thread-2406001-6-1.html	[03-06] «通卡小助手(日日更新) - 警岛民生 - 厦门小鱼...
BBS_和迅论坛	4	bbs.online.sh.cn/forum/forum-1192-1.html	热点爆料 - 上海热线互动社区
BBS_湖北东湖社区	1	bbs.yynet.com/viewthread.php?tid=808865&extra=page%K3D1	您认为房价上涨有哪些原因?
BBS_天涯论坛	7	www.tianya.cn/publicforum/articleslist/0/free.shtml	天涯杂谈_天涯论坛_天涯社区
BBS_天涯论坛	4	www.tianya.cn/new/publicforum/articleslist.asp?pageNO...	天涯杂谈_天涯论坛_天涯社区
BBS_中文之家论坛	1	bbs.yynet.com/viewthread.php?tid=861548&extra=page%K3D1	天涯杂谈_天涯论坛_天涯社区
BBS_深圳之窗社区	3	bbs.ngzb.com.cn/viewthread.php?tid=74832	最近官员语录大比拼, 哪条最雷最牛?
BBS_南江社区	1	www.tianya.cn/new/publicforum/articleslist.asp?pageNO...	天涯杂谈_天涯论坛_天涯社区
BBS_雷傲论坛	9	bbs.yynet.com/viewthread.php?tid=1021218&extra=page%K3D1	天涯杂谈_天涯论坛_天涯社区
BBS_杭州网络论坛	1	bbs.yynet.com/viewthread.php?tid=1021218&extra=page%K3D1	天涯杂谈_天涯论坛_天涯社区
BBS_19楼论坛	10	www.tianya.cn/new/publicforum/articleslist.asp?pageNO...	天涯杂谈_天涯论坛_天涯社区
BBS_论坛	0	bbs.zmfish.com/thread-2406001-5-1.html	[03-06] «通卡小助手(日日更新) - 警岛民生 - 厦门小鱼...
BBS_论坛	11	www.tianya.cn/new/publicforum/articleslist.asp?pageNO...	天涯杂谈_天涯论坛_天涯社区
BBS_C51技术社区	1	bbs.yynet.com/viewthread.php?tid=848573&extra=page%K3D1	关于《彩票就是赌博》一文向所有网友的补充说明
BBS_DnewsIT写作	3	bbs.ngzb.com.cn/viewthread.php?tid=133735	早报网联合广西邮政征集贺卡DIY秀、新年祝福语、大奖等等...
BBS_爱车汽车论坛	0	www.tianya.cn/publicforum/articleslist/0/worldlook.shtml	国际观察_天涯论坛_天涯社区
BBS_奥一论坛	1	www.tianya.cn/new/publicforum/articleslist.asp?pageNO...	国际观察_天涯论坛_天涯社区
BBS_大洋论坛	8	www.tianya.cn/publicforum/articleslist/0/new.shtml	天涯时空_天涯论坛_天涯社区
BBS_网络论坛	0	www.tianya.cn/publicforum/articleslist/0/new.shtml	天涯时空_天涯论坛_天涯社区

线程	编号	URL名称	URL标题
News_新浪新闻 - 生活	0	blog.39.net/sinderer/a_3446800.html	女人给男人的4个忠告
News_搜狐新闻 - 国内新闻	4	blog.39.net/foshanzys/a_3717761.html	宫颈癌患者的生育救星
News_搜狐新闻 - 国内新闻	0	blog.39.net/zgyzm/a_3632772.html	营养师原创: 吃青菜剩下的菜汤要不要喝?
News_搜狐新闻 - 财经中心	4	blog.39.net/foshanzys/a_3739917.html	谁下了毒手, 让女性患上输卵管
News_搜狐新闻 - 体育	0	blog.39.net/dingyisheng169/a_3434763.html	谢光志培需要关注什么问题
News_搜狐新闻 - 体育	4	blog.39.net/gzra/a_3755705.html	男性不育症分为哪几类
News_搜狐新闻 - 教育	0	blog.39.net/dingyisheng169/a_3314737.html	最具颜值新标准出台!
News_搜狐新闻 - 文化	2	gb.cri.cn/20864/2010/01/25/1945s2741036.htm	贝嫂时尚秀秀长腿 - 与猛男拥抱激吻脸(图)
News_中国青年报 - 新闻中心	0	blog.39.net/songqingjies/a_3458445.html	最靓草幕斯蛋糕
News_中国青年报 - 教育	0	jbk.39.net/keshi/neike/shen/4d229.html	慢性肾炎
News_中国青年报 - 教育	3	blog.39.net/hymjiaoshou/a_2315831.html	癌症不用治 会自行消失?
News_中国青年报 - 教育	0	jbk.39.net/ks/fuchanke.html	妇科疾病
News_中国青年报 - 教育	0	jbk.39.net/keshi/neike/neifenni/49098.html	高血脂
News_中国青年报 - 教育	0	bbs.39.net/iframe/1jforjkb-319273.html	慢性肾炎
News_中国青年报 - 教育	2	gb.cri.cn/20864/2010/01/25/1945s2741068.htm	组图: 史上最美女体育女主播靓丽写真
News_中国青年报 - 教育	0	www.39.net/gp/wydrk/index2.html	慢性肾炎
News_中国青年报 - 教育	0	bbs.39.net/iframe/1jforjkb-299160.html	高血脂
News_中国青年报 - 教育	1	blog.39.net/foshanzys/a_2427168.html	实习生只不过是“潜规则”
News_中国青年报 - 教育	0	blog.39.net/shanquinn/a_1229370.html	单挂阿阿阿的艾类博客分类目录链接
News_中国青年报 - 教育	1	blog.39.net/huanglijuan/a_1311997.html	一位医生的愤怒: 《说说医院医生护士之间的潜规则》背后的...
News_中国青年报 - 教育	0	blog.39.net/buyun120/a_1647752.html	正常卵泡大小是多少?
News_中国青年报 - 教育	2	gb.cri.cn/20864/2010/01/25/1945s2741104.htm	拉姆世界杯后完婚 欲三天内完成人生两大梦想
News_中国青年报 - 教育	4	blog.39.net/BabyCat0716/a_3708462.html	你打算怎么吃?
News_中国青年报 - 教育	4	blog.39.net/dai19830701/a_3750645.html	2010年1月4日
News_中国青年报 - 教育	4	blog.39.net/XEPN8NC3K3C3K7NE9K9K9EE/a_3791805.html	要BB的冲动竟然来自八
News_中国青年报 - 教育	2	blog.rol.com.cn/1427/article_1426308.html	联通iPhone先策, 会采取补救措施
News_中国青年报 - 教育	0	blog.rol.com.cn/1550/article_1549821.html	转译搞笑笑: 2010年是新I3ISIT处理器特点和区别
News_中国青年报 - 教育	3	blog.39.net/wyssh/a_2164474.html	癌症疼痛并非“终末期”来临
News_中国青年报 - 教育	3	blog.39.net/hymjiaoshou/a_2019075.html	内向的人更容易患癌症
News_中国青年报 - 教育	3	blog.39.net/zmiil/a_2020292.html	怎样预防癌症疼痛的产生
News_中国青年报 - 教育	3	blog.39.net/hymjiaoshou/a_2108171.html	能越睡痛痛复“三进关”
News_中国青年报 - 教育	4	blog.rol.com.cn/1338/article_1336689.html	湛江风城(三)
News_中国青年报 - 教育	3	blog.39.net/yunqihethly/a_2007587.html	李朝博原病的饮食对策
News_中国青年报 - 教育	3	blog.39.net/xiaobingbu/a_1871267.html	莲藕 糖尿病患者最好的食物
News_中国青年报 - 教育	1	blog.39.net/tianfangwen/a_1951332.html	这位男人该如何处置
News_中国青年报 - 教育	1	blog.39.net/KETN8N83KA4KCN8K8N8K4K82KDDKCC3C3/a_20...	范长伟: 我为糖尿病的灸法

舆情采集特色功能介绍如下：

3.1.1 支持网页结构自动分析

论坛和新闻评论的特点是一个物理网页中一般包含多个贴子（或评论），而且随着时间的推移，同一个网页中的贴子会动态变化，有新的子贴增加，也可能有旧贴删除。搜索引擎的处理方法是一个网页解析出一份内容，这样的处理方式不能满足舆情监控的需求，对于舆情系统来说，每一个贴子都代表着一份有价值的信息，因此一个论坛或评论网页往往需要分解为 N 个贴子，每个贴子都作为一条独立的信息参与舆情分析。

eOMS 舆情采集系统实现了基于网页结构的统计分析算法，能够自动识别、分隔出论坛和新闻评论中的每一个贴子及其评论信息，并通过 URL 特征分析和模板识别技术实现主贴和跟帖的区分，并可以识别跟帖所属的主贴，对于跟帖分页等多种情况都可以进行正确处理。通过正确判断主贴和跟帖为下一步实现舆情分析打下了基础。

3.1.2 支持基于模板的元数据解析

在贴子自动分隔的基础上，eOMS 舆情采集系统支持基于模板的元数据解析功能，对于每个贴子可以解析出发贴人、发帖时间、贴子主题、贴子内容、点击数等元数据属性，这些属性是舆情分析的重要数据。

3.1.3 内嵌 Javascript 脚本执行引擎

随着 Web 2.0 相关技术的发展，脚本语言越来越多地应用于论坛、新闻评论、博客等类型网站的建设，目的是提高用户的交互性和易用性。由于搜索引擎的采集工具一般不支持脚本语言，因此 Google、百度等搜索引擎只能收录少量来自论坛、博客和新闻评论的内容。

eOMS 舆情采集系统内嵌了脚本引擎，能够实现网页中脚本语言的自动解析和执行，从而能够实现对于采用脚本语言的论坛、博客以及新闻评论网站的采集，为舆情分析提供全面的数据保障。

3.1.4 支持 RSS 解析

许多博客/新闻评论采用 RSS (Really Simple Syndication 或 RDF Site Summary, 信息聚合技术) 实现内容共享, 除了网页的基本信息, RSS 一般都会包含更丰富的元数据信息, 比如标题、作者、发表日期、分类、关键词等, 有些 RSS 甚至包含更多的有价值的信息。这些元数据对于舆情分析往往具有重要的价值, eOMS 舆情采集系统实现了对 RSS 聚合内容的自动解析和采集。

3.1.5 支持境外网站监控

系统集成代理技术, 通过代理服务器, 绕过我国出、入口防火墙的屏蔽, 浏览被屏蔽境外网站, 监控有关涉及本地的信息。

eOMS 舆情采集系统在支持通用网页采集功能的基础上, 实现了上述特色功能, 从而能够实现对论坛、博客、新闻评论等内容的及时精确采集, 为舆情分析提供了丰富的数据资源。

3.2 舆情信息存储和智能检索

从互联网搜索到的舆情信息存储在 eOMS 舆情监控系统内置的全文数据库系统中。全文数据库的优势是存储海量的非机构化的信息, 其采用实时动态索引技术, 数据增删改时快速同步更新索引, 无需重建整个索引以及局部重建索引, 即数据变化后马上能检索出来, 保证了信息搜索的实时性和有效性, 满足舆情应用的核心检索需求。依靠专业的全文数据库存储系统, eOMS 舆情监控系统可以对大量监控数据源进行长期监控, 存储长达数年的数据, 方便我们对数据进行分析、比较、统计。eOMS 舆情监控系统的智能检索在性能、查全、查准、索引膨胀率等指标上都表现优异。

3.2.1 中文分词

中文分词是全文检索需要的技术之一, 进行分词后能使全文检索结果更加符合用户习惯, 更加接近用户的期望结果。中文分词后台有词库作为支撑, 用

户也可以根据自己的需要和行业特色来添加和维护词库。我们的中文分词在准确率和召回率指标上都有很好的表现。在中文分词的技术上，通过关键词空间向量提取、主题识别、权重组合等技术，达到热门主题识别的目的。

3.2.2 内容相似性分析

eOMS 舆情监控系统基于内容相似性算法，找出内容相似的文章并根据相似度排序，包括做相似性关联标识、相似性去重等处理。

3.2.3 基于规则自动分类

该功能可以根据用户定义的关键词词典和关键词规则，遍历扫描文章内容，识别出关键词并标识这些文章，可以实现对包含关键词的文章的自动分类标识。

3.2.4 基于内容自动分类

该工具根据样本训练得出分类特征向量空间模型、然后根据文章的特征向量实现对文章的自动分类标识。

3.3 舆情管理

3.3.1 舆情信息分类

eOMS 舆情监控系统可以针对不同的用户或角色定义和维护舆情分类体系，系统支持多维度、多层分类体系规划，用户可以采用多种灵活的方法定义和维护分类体系，维护分类检索规则。

3.3.2 舆情用户、组织管理

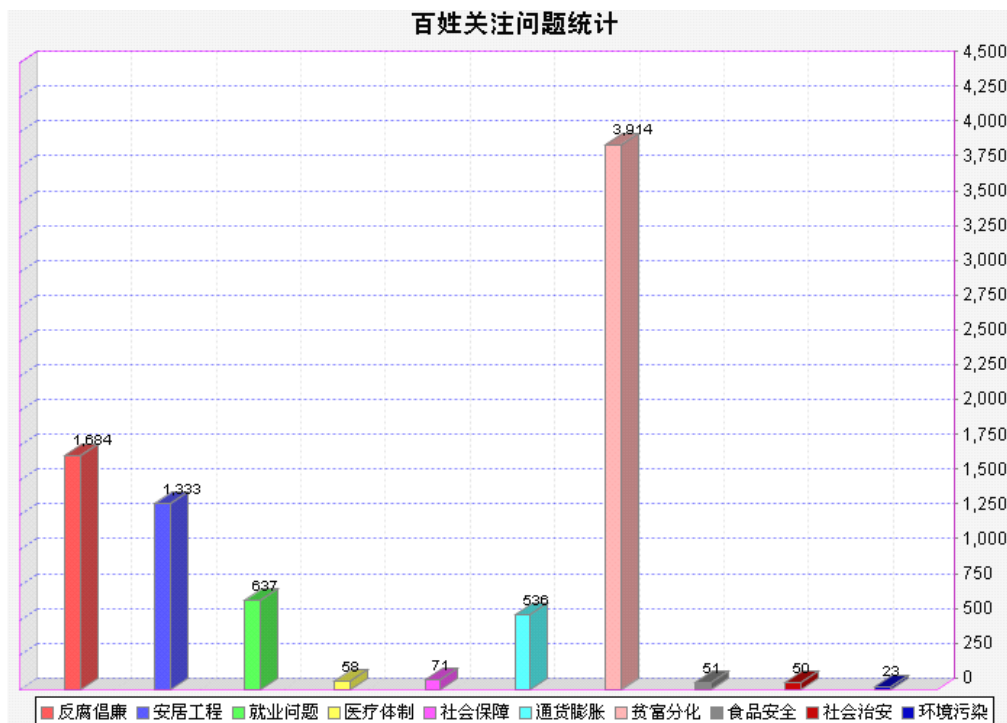
支持舆情组织体系规划：系统支持权限、角色和用户“三要素”应用模型，用户可以根据组织的需要，灵活的建立工作角色（管理员、舆情采集人员、舆情加工人员、舆情浏览人员），并设定用户和角色的对应关系。

3.3.3 舆情统计、分析

eOMS 舆情监控系统提供灵活的统计分析工具，可以按照用户的需要构建个性化的统计分析图表。舆情统计分析工具可用于构建：

(1) 舆情趋势分析与预警

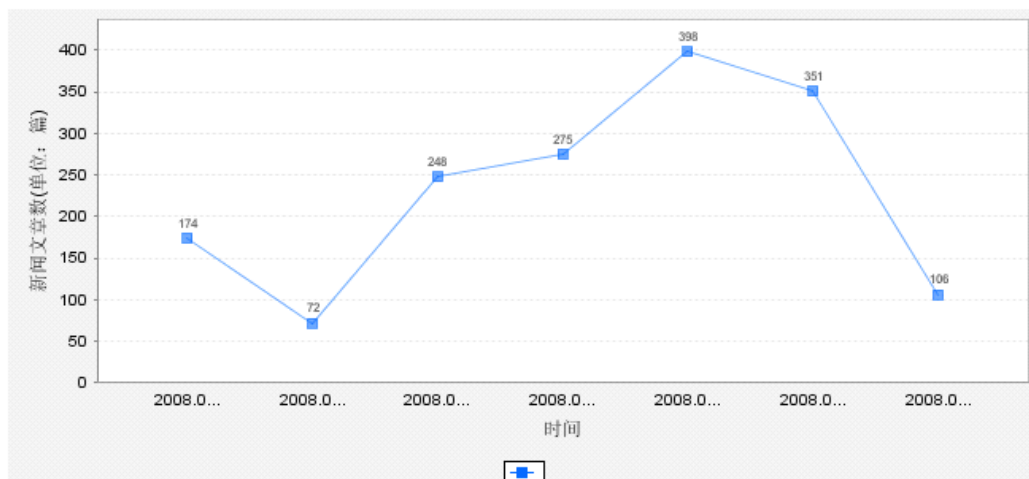
在舆情分类和褒贬分析基础上，建立多个舆情指标，可以根据舆情指标和舆情分类的结果评估和分析当前的宏观舆情态势，自动给出预警信号，辅助相关人员对舆情进行干预和引导。另外，还可提供全国与各地舆情对比功能，帮助用户及时准确地把握宏观舆情走势。



(2) 热点事件追踪

对于热点事件，eOMS 舆情监控系统可以做到持续追踪，并通过趋势分析图等技术帮助用户了解热点事件的报道趋势，帮助用户更好地对舆情进行研判。

热点专题西藏暴乱最近几天趋势图：



时间列表：

1	2008.04.05	174
2	2008.04.06	72
3	2008.04.07	248

(3) 舆情观点归纳

对于大多数论坛和博客，主贴往往是一个话题和观点的发起者，而跟帖才是网民对某一问题和事件态度的表达，eOMS 通过对某一主贴跟帖的统计分析，归纳网民的主要观点，并统计网民的褒贬倾向分布情况，帮助用户更准确地了解社情民意。

3.4 舆情服务

通过 eOMS 舆情监控系统可以建立舆情协作办公平台：

全部 博客 论坛 政府媒体

搜索
关键词: 反腐倡廉 通货膨胀 贫富分化 社会保障

分类树

- 系统分类
- 百姓关注
 - 反腐倡廉
 - 安居工程
 - 就业问题
 - 医疗体制
 - 社会保障
 - 通货膨胀
 - 贫富分化
 - 食品安全
 - 社会治安
 - 环境污染
- 地方舆情
 - 北京市
 - 上海市
 - 天津市
 - 重庆市
 - 河北省
 - 山西省
 - 辽宁省
 - 吉林省
 - 浙江省
 - 黑龙江省
 - 江苏省

反腐倡廉 更多 >>

[原创]咸宁中院、司法腐败到了让人不能忍受的地步!

文章来源: 湖北东湖社区 更新时间: 2010.01.27 14:23:44

警方认定南勇牵涉反腐窝案 据称受贿数十万元

文章来源: ZOL博客 更新时间: 2010.01.23 18:58:55

中国的铁饭碗就是腐败

文章来源: ZOL博客 更新时间: 2010.01.23 12:03:11

安居工程 更多 >>

海南房价再度陷入疯狂 外地老板一伙买下60套房

文章来源: 搜狐社区 更新时间: 2010.01.27 14:14:53

高房价、GDP数据：中央、地方、居民的博弈何时休!

文章来源: 天涯 更新时间: 2010.01.27 13:44:42

取消预售房制度只能助涨房价

文章来源: 中金博客 更新时间: 2010.01.27 13:01:28

舆情区域分布

■北京 ■天津 ■上海 ■重庆
■河北 ■河南

百姓关注问题统计

■反腐倡廉 ■安居工程
■就业问题 ■医疗体制
■社会保障 ■通货膨胀

(1) 舆情部门搭建门户：利用舆情监控系统可便捷地搭建出供最终舆情使用者访问的舆情门户，不同的用户登录后访问的内容有所不同。增强了最终用户的访问体验，提高了舆情利用的效率。

(2) 舆情门户、分类导航和舆情推荐：舆情工作部门可利用舆情监控系统建立舆情信息门户为舆情服务对象提供舆情服务，也可以通过舆情分类体系提供舆情分类导航服务。

(3) 检索功能：舆情监控系统提供对舆情信息的全文检索、标题检索、来源检索、日期检索等功能，并且支持智能检索以及相似检索。

(4) 个性化定制：不同访问用户可以定义不同的首页，并根据自己喜好配置首页的显示内容、样式、摆放的位置、信息条数、统计分析内容，方便用户对系统的使用。

(5) 邮件订阅：用户可以根据要求，分类定制舆情信息，收藏重要的舆情信息，以及通过订阅获取舆情邮件自动送服务，实现个性化和及时便捷的舆情服务。