

IFACE 专业搜索 2.1

联索 IFACE 可以定制信息源，对搜索引擎进行精确导航，对信息内容进行标记，自动获取并管理海量信息中的精华，在经过智能化信息处理的专业存储库中随意搜索和呈现个性化内容。

本简明指南包括如下内容：

1. 概览
2. 按专题的形式组织信息
3. 选配信息源
4. 自动分类和整理信息
5. 标注摘录内容
6. 管理搜索
7. 使用越多，结果越好

1. 概览

联索 IFACE 可以从成千上万的信息源中自动获取精华信息，利用精确制导的网络蜘蛛，聚合、过滤、摘录、整合信息，通过自定义分类，自定义数据属性项，即可自动采集并保存网页、文档、图形、表格、影音等格式文件，还能生成结构化的专业数据库，对海量数据实现精准检索，是网站、企事业和政府实现门户搜索、电子商务、数字媒体、信息资源管理、信息整合增值、公众信息服务等应用的理想系统。

用户登录

专业搜索管理系统支持多用户同时使用，并可灵活控制用户的功能权限及操作对象。为此，您需要注册一个用户帐号，或者咨询系统管理员。

进入搜索管理系统后，您可以方便地在用户首页使用各种常用功能。

A 自定义搜集专题

将要搜集的信息分门别类放在不同专题中。例如，财经、资讯、娱乐等，您可以根据需要自由定义。

B 管理信息源

精确制导，从众多的信息源中自动获得精华信息，聚合、过滤、摘录、整合……

C 自定义分类

自定义多种分类标准，信息搜集的同时即可对文档自动分类。

D 网页摘录

用户网摘融合搜索引擎的强大功能，机器批量采集配合人工精挑细选，增强个性化搜索功能。

E 内容管理

提供内容管理、信息发布的基础支持

F 搜索

提供可扩展的企业级搜索功能，帮您定制自己的搜索引擎。

G 服务控制

设置和控制专业搜索的各项系统核心服务。

2. 按专题的形式组织信息

专题是特定主题信息或者行业信息的目录组织方式，使得系统自动搜集的信息可以按照类别进行划分，专题可以参照标准分类体系、用户使用习惯、或者业务需要进行灵活定义。

A 专题搜集

利用专题方便地规划搜索引擎的信息库

B 专题列表

内容组织及各专题的主要信息一目了然。

C 保存方式

收录文档以数据库或者文件等不同形式保存。

D 全文检索和自动分类

可指定各专题是否启用全文检索或自动分类功能。

E 搜集时效

可以控制各专题的搜集时效，包括文章时限及归档有效期等。

F 创建及修改时间

系统自动记录对专题的操作时间。

G 信息源数

系统自动计算各专题的信息源数。

H 文档数

系统显示各专题下的收录文档数。

I 就绪状态

显示各专题的操作状态，就绪的专题即可开启信息搜集。

3. 选配信息源

信息源是信息来源的简称。信息源的类型可分为网站、数据库、文件系统目录等，信息可以来自外部的互联网或者内部的企业网。引入信息源的精确控制机制，可以有效提高信息处理的精深程度和效率。例如，网站信息

源可以是目标网站范围、网站频道范围、网页范围、网页内任意信息块（数据区）范围直至属性值对等，信息源同样可以按照类别进行组织和管理。

信息源的配置

灵活的人工设置和自动设置，调控信息搜集的广度和精度。

A 专题列表

您可以在此快速定位和选择各专题下的信息源。

B 信息源列表

列出当前专题下各信息源的主要信息，点击信息源名称即可进行设置。

C 设置信息源

通过设置信息源，您可以从多方面精细控制信息的搜集与处理精度，同时获得更加准确的搜索结果。

D 网站收藏夹

专门提供互联网信息源信息搜集的管理。

E 信息源状态

自动追踪网站信息源的设置状态。

F 信息源统计

自动统计网站信息源的分布情况。

4. 自动分类和整理信息

利用计算机程序对海量的文档进行分类，将文档划归到自定义类别体系（或主题）下的某一具体类别中。虽然目前自动分类并不能完全代替人工所做的相关工作，但是提供了一个最大程度降低成本和人工工作量的方法。

通过自定义分类标准，系统可以完成多维度的自动分类。除了导入已有的分类标准，您也可以随时对类别进行定制，并根据需要加入特征词。

在信息搜集处理过程中即可对文章进行自动分类，还可以对信息进行结构化处理，提取元数据，提取数据区的属性数据、转存到数据库中，这样，您就可以轻松建立起专业信息库（数据库）。

一般网页都有不同的版面区域来放置不同的内容，系统通过对网页划分区域，进行分割，实现对多媒体信息进行结构良好的组织，也根据需要过滤掉信息。

在信息块中，可以进行细致的数据提取，系统剔除不相关信息和提取数据对象的属性及其内容。例如：在数据区中，提取有关数据对象的信息，如产品价格、产品型号等。

信息提取技术能够最大程度地发挥电子商务应用的潜力。

说明：网站类型的信息源，内容页中充斥着大量的无关信息，例如，广告条、广告链接、广告图片、产品目录、导航条、版权声明、用户反馈等。真正有价值的内容，如标题、正文、时间等。系统能够甄别有用信息，过滤掉冗余的、无用的或者不相关的信息甚至垃圾信息。

5. 标注摘录内容

网页摘录，又称网摘，满足用户在线研究的需要。帮助用户组织网络内容，积累信息知识，具有对内容页的收藏、分类、摘录、加注标签、保存到信息库、信息库共享等功能。

A 个人网摘

收藏、整理用户个人感兴趣的文章，以后就可利用强大的搜索引擎快速定位它们，免去大海捞针的烦恼。

B 标签

您可以给收藏的网摘加标签，系统会自动为您管理与统计。

C 编辑

可以对自己私有的内容进行编辑修改，以符合自己的需要。

搜索关键词不仅可以是文章中出现的文字，还可以是您自己加注的标签。

6. 管理搜索

搜索管理系统是对专题、信息源，信息库、系统配置、专业词库等，以及搜索系统的运营维护进行管理的系统。系统的主要用户是行业专家、搜索编辑（主编或栏目编辑）、系统管理员和相关业务人员。

A 内容管理

搜集及提供高质量的信息内容，供用户查询和利用。

B 词库管理

管理词库和自动分类的规范和标准等中文信息处理的语言资源。

C 用户管理

管理用户及其功能权限。

系统管理

利用系统管理功能，可以对搜索引擎系统的核心服务进行配置与管理。这些服务包括信息搜集服务、信息索引服务以及信息检索服务等。这部分功能的主要用户是系统管理员或系统维护人员。

A 服务列表

在此选择专业搜索的各项核心服务，来查看相关信息。

B 服务状态

显示专业搜索各项服务的运行状态。

查询的图示,系统的幕后工作,其实还有更多……

A 检索服务器

指定搜索要使用的检索服务器。

B 关键词

指定搜索关键词,支持多关键词或整句。

C 时间

可以指定搜索结果的时间范围。

D 文字位置

可以对文章的不同位置进行搜索,例如标题、正文。若是搜索用户网摘,还可以对文章标签进行搜索。

E 专题范围

可限定在特定的专题内进行搜索。可以指定一个或多个专题。

F 搜索结果

搜索结果中展示了命中文章的主要信息,包括标题、时间、摘要、文章地址等。另外,开发者或集成商在应用系统中还可以利用丰富的接口,定制各种其它的搜索结果展现形式。

C 其它信息

显示各服务的相应信息,例如信息搜集进度,索引的建立情况,检索服务的负载等。

D 选择服务器

管理器可以控制多台服务器上的服务,不论是本地还是远程。

E 服务启停

启动或停止专业搜索的后台服务。

F 服务设置

可以在此指定专业搜索服务的各项设置,包括服务线程、索引更新间隔、检索服务器更新目标等。

7. 使用越多,结果越好

将用户的查询输入关键词,和用户点击的搜索结果相关联,结合语义分析算法以及用户行为分析的算法,根据数据统计分析,自动淘汰低质量信息内容,始终维护一个高质量的信息库,使用时间越长、次数越多,信息库质量就越好,搜索结果也越准确!下面仅仅是一个搜索