# 海洋空间信息分布式数据挖掘平台（含软硬件）

## 一、系统技术要求

* **技术规格要求**

## 技术要求

1. 软件基于组件技术开发，支持Web Service开发接口；
2. 平台服务器集群采用Ubuntu Server 12 及Windows Server 2003以上等支持多用户、多任务的操作系统，结构化数据可选择部署在SQL Server 2005及以上等数据库中；非结构化数据部署于Couchbase、MongoDB等高性能开源分布式NoSQL数据库
3. 应用系统的开发及运行结构是基于后台数据库的三层架构，即Web服务器、应用服务器和数据库服务器；

## 平台软件需求与开发内容包括：

1. 能快速处理TB级别的结构化及非结构化数据
2. 支持对多个网络爬虫的并行控制与多个网络爬虫数据包的读取
3. 支持对TB级WEB数据的快速降噪
4. 支持对TB级网页js代码的分布式解析
5. 支持对TB级网页的定制式数据智能提取
6. 支持对TB级内容近似数据的快速智能去重

## 平台集群工作站基本技术要求

1. 品牌要求：知名品牌
2. 机架式服务器，单台配置要求：2颗E5-2630 以上, 4 x 16G内存以上，3x600G 3.5英寸 15000转 SAS硬盘以上，R500,RAID5, 1个800W电源
3. 服务要求：3年上门

## 二、项目实施与售后服务要求

## 技术支持与售后服务承诺

1. 供货单位应提供三年的免费的软硬件技术支持和服务，承诺三年软件免费维护。
2. 如遇有二次开发工作，涉及底层结构变动，可根据双方协商收取一定费用。
3. 本项目所要求的技术服务包含但不限定于上述技术服务项目，供货方可以根据自己的服务经验及能力提出更多的适合于本项目的技术服务内容。

## 培训要求

1. 供货单位必须根据要求提出全面的培训计划，并在合同签订后实施。
2. 供货人负责所有培训费用（含培训教材费）及各项支出包括项目单位所在地城市与培训地之间的往返交通费用，当地的交通费用，食宿费用，每人日培训费用等。

***注: 供货单位必须对上述的所有内容进行整体报价和供货,不得拆分***