备用话术：

随着WEB技术与5G技术的迅猛发展，利用互联网的云端存储也成为了现今主流互联网企业的数据存储方式，这源于互联网数据的爆发式增长，每一天都会有成千上万的用户数据来往于数据库与用户终端之中。海洋数据虽不比互联网数据的

而如何将这些丰富多样的数据以低门槛的要求，以获取的形式共享给公众

目录

[基于多源海洋数据的信息平台开发与应用研究 4](#_Toc6220855)

[摘要： 4](#_Toc6220856)

[Development and Application of Information Platform Based on Multi-source Ocean Data 5](#_Toc6220857)

[Abstract： 5](#_Toc6220858)

[1 绪论 6](#_Toc6220859)

[1.1研究背景 7](#_Toc6220860)

[1.2研究的目的与意义 8](#_Toc6220861)

[1.3 国内外研究现状 9](#_Toc6220862)

[1.4 研究内容与章节安排 11](#_Toc6220863)

[1.4.1 研究内容 11](#_Toc6220864)

[1.4.3章节安排 12](#_Toc6220865)

[2 系统相关技术研究 12](#_Toc6220866)

[2.1Bootstrap前端框架 13](#_Toc6220867)

[2.2 Web服务器网关接口 14](#_Toc6220868)

[2.3 B/S架构模型 15](#_Toc6220869)

[2.3.1 B/S架构模型与C/S架构模型的介绍与比较 15](#_Toc6220870)

[2.3.2B/S结构的特点和优势 16](#_Toc6220871)

[2.4MySQL数据库 16](#_Toc6220872)

[2.4.1 MySQL数据库介绍 16](#_Toc6220873)

[2.4.2 MySQL数据库所需关键知识 17](#_Toc6220874)

[2.5 Python-flask web框架及关键模块 18](#_Toc6220875)

[2.5.1 Python-flask框架介绍 18](#_Toc6220876)

[2.5.2 Flask-Uploads模块 19](#_Toc6220877)

[2.5.3Flask-SQLAlchemy扩展 19](#_Toc6220878)

[3 需求分析及可行性分析 22](#_Toc6220879)

[3.1系统需求分析 22](#_Toc6220880)

[3.2平台需求概述 23](#_Toc6220881)

[3.2.1业务功能需求 23](#_Toc6220882)

[3.2.2 平台性能需求 24](#_Toc6220883)

[⑴响应时间指标 24](#_Toc6220884)

[⑵系统容量性能指标 24](#_Toc6220885)

[3.3 平台用例分析 24](#_Toc6220886)

[3.3.1 海洋数据查看与下载 25](#_Toc6220887)

[3.4 平台可行性分析 25](#_Toc6220888)

[3.4.1技术可行性 26](#_Toc6220889)

[3.4.2经济可行性 27](#_Toc6220890)

[4 平台总体设计 27](#_Toc6220891)

[4.1 平台架构设计 27](#_Toc6220892)

[4.2平台的框架设计 28](#_Toc6220893)

[4.3 平台的模块设计 29](#_Toc6220894)

[4.4 平台数据库设计 30](#_Toc6220895)

[4.4.1 实体描述 30](#_Toc6220896)

[4.4.2 数据库迁移 31](#_Toc6220897)

[4.4.3 数据库表格设计 31](#_Toc6220898)

[5 基于多源海洋数据信息平台的实现 36](#_Toc6220899)

[5.1 项目前期工作 36](#_Toc6220900)

# 基于多源海洋数据的信息平台开发与应用研究

# 摘要：

科教兴国，海洋强国。中国是一个正在高速发展的沿海大国，具有丰富海洋资源和海洋科学数据，作为国家基础战略资源，是认识海洋，发展海洋的重要战略前提。但是海洋数据丰富和繁杂，公众如何根据自己所需要的信息从海量的数据文件中获取有效的数据文件，以及如何保证信息的实时性，成了当今海洋数据共享的一大重点问题。

论文围绕着多源海洋数据的信息平台的开发，首先分析了B/S与C/S架构对平台建设的基础性方向的决定性、不同开发语言对于所选择架构的支持性、开发语言的已有框架等平台开发的相关技术，同时对平台相关流程及数据库建设进行了具体研究。

在确定平台架构，开发语言，所选框架的基础上，选取了最为合适的主流数据库，并根据海洋数据的时效性和多样性等特点，采取了最为合理的数据组织与管理方式，建立了符合当前海洋数据文件的数据库；设计并实现了海洋数据分类展示，海洋数据分类查找，时效性更新，海洋数据文件及时下载等功能。

结合实际的现有海洋数据文件，根据不同类型分类展示，对海洋数据关键要素进行了直接展示，方便公众获取直接有效地海洋数据，主要工作如下：

1.采用B/S三层架构模型，以网站的形式，降低了海洋数据的获取门槛，使大量基础数据向公众进行开放。

2.采用MySQL将海量的数据文件存储在云端，保障数据的安全性与及时的有效性，不同类别的海洋数据以模型的方式，利用Python-flask框架映射至MySQL数据，各类海洋数据文件独立性强，耦合性低，方便用户下载。

关键词：海洋数据 信息平台 Python B/S架构 Python-flask框架

# Development and Application of Information Platform Based on Multi-source Ocean Data

# Abstract：

**Key words:** **Ocean Data;Information platform;B/S Framework;Python-flask Frame;Python**

# 1 绪论

随着我国国力的不断发展进步，海洋的开发与保护也得到了高度的重视，海洋强国随之成为了国家的发展战略，作为一个拥有着漫长的海岸线的过渡，这也注定着我们可以获取丰富的各种类别的海洋数据，例如：海洋生物，海洋水文，海洋地形等数据，不同的数据都拥有着各自应用的前景，海洋生物用于海生养殖及珍稀动物保护，海洋地形于航海研究等。要建设海洋强国就必然要全民参与，而如何将研究机构，船舶测量数据通过简易，低门槛的方式共享给社会各界有需要的组织与机构，就成为了现在需要解决的重要问题。

## 1.1研究背景

进入新世纪以来，全球社会与政府越来越重视海洋领域的探索与研究，并向海洋数据等领域投入了大量的资金与人力。随着互联网技术的快速发展，海洋勘探领域技术的飞跃进步，海洋数据已经进入了爆炸发展的时期。海洋数据的处理与共享也成为了互联网技术的最重要的应用场景，海洋数据的‘海量性’、‘多类性’、‘模糊性’、‘时空过程性’‘动态更新频繁’已经成为了其最具特色的价值所在，收集整理，分析研究这些信息的经济及生态价值也成为了现今各个研究机构的一大课题，由此也引出了一个建设的大方向——海洋信息化。

海洋信息从根本上分为两大类，一类是海洋自然科学类，一类是海洋社会科学类。海洋自然科学类主要指的是运用各种技术手段和设备进行调查、探测、分析而得到的信息，其中使用得到的海洋设备包括卫星、飞机、探测船、浮标、陆地/海洋站。包括了以下几种内容，海洋化学（水体温度、pH值、溶解氧、电导率等水质参数的检测），海洋生物（通过全球各国海洋部门数据集与各研究机构数据集，综合了浮游生物、浮游植物、初级生产力、鱼类等数据），海洋水文（通过海洋台站、观测浮标、调查船等观测手段获取，包括温度、盐度、波浪及海流等要素）、海洋气象（通过固定陆地站、海洋站、观测船、观测浮标等观测手段获取，包括海面气温、气压风向、风速、风向、位势高度等要素）。又因为原始的海洋数据资料不能直接用来展示与使用，因此需要预先对数据进行清洗、转换、选择等操作，最终展示在平台，供研究机构下载与使用。

海洋社会科学类是与海洋自然科学类相比较而言的，截至目前还没有研究机构或者政府对此有明确的定义。从目前的发展情况来看，我国的海洋强国战略，海上丝绸之路等人文经济领域的相关信息，都可以划分到这一类信息当中，大致可以划分为海洋战略信息，海洋人文信息，海洋经济信息三大类。海洋战略信息指的是从长远角度来看，我国实现真正的海洋强国目标，政府智库，官方媒体，政府海事单位等相关部门对整个国家海洋发展的规划，以及面对与他国海洋发生主权纠纷，海洋资源合作开发等长期性发展问题所做出的科学的考量与决策。

海洋人文信息宽泛定义为在人类与海洋的长期接触与活动的过程中，形成的精神与文化现象。包括文化，政治，思想等方面，例如“郑和下西洋”所产生的强大祖国荣耀感对于实现海洋强国战略是必不可少的精神支柱，只有形成了宏观的海洋意识，具有宽广的海洋人文知识素养与国际海洋视野，才能培养出真正热爱海洋事业，甘愿贡献自身于海洋事业的栋梁之才。

海洋经济数据主要指的是在开发海洋资源产生经济利益而进行生产活动的过程中，产生的相关数据，这类数据在引导企业进行合理的生产方面起到了巨大的作用，帮助企业收集各类有意义的信息，通过分析处理，根据企业的运营与业务需求，做出有益于企业长足发展的科学可持续化决策，拥有产生高效益的能力，科学的海洋数据也有助于减少企业的运营成本，从数据中求高收益，从科学决策中获取商业优势，成为海洋领域的优秀企业。海洋经济领域从海洋数据中得到更好地发展，其所产生的经济价值也会反过来促进海洋数据信息的收集与研究，帮助研究机构获得更多的资金与人力支持，对于探索海洋数据有百利而无一害，如此而形成的良性循环，有助于我国的海洋强国战略的实施。

## 1.2研究的目的与意义

“数字海洋”的诞生是源于“数字地球”的建设，二者都是为了将地球上所有探测到的数据以数据库的形式存储，并共享给全球社会，集全球的科研精力及资源，挖掘有助于人类可持续发展的有用信息，助力社会发展，而作为一个海洋面积约占71%的星球，“数字海洋”也必然成为了这一庞大建设的最主要部分。

基于海洋数据的信息平台的建设，可以将丰富海量的海洋数据信息，直观便捷的展示给公众，按照不同的海洋数据分类以基于互联网网站的形式，相较于无分类的海量的数据集文件，网站可以让用户根据自己的使用需求选择来自不同海域的数据集以及在线下载，以用来将来进一步的分析与研究。友好的界面交互与UI设计有助于激发公众对于海洋知识的热爱，有助于海洋知识的基础科普，在展示过程中采用数据集表格的形式，用户可以看到自己所选择数据集的所有数据，包括数据格式，数据上传时间，更新频率等相关字段，保证了公众获取数据的及时性与可用性。

基于海洋数据的信息平台的建设，完成了从庞大数据库到可视化信息查看的转变，用户不需要经过相关的专业训练，只需要基本的操作就可以查看并下载海洋生物、海洋化学、海洋水文等信息。从政府角度来看，相关研究机构只需将自己机构的研究数据与结果上传至海洋数据信息平台，政府相关部门便可以根据当前决策需要，搜索需要的海洋数据，结合数据进行科学决策，去除了政府相关部门与研究机构的沟通隔阂，降低了信息共享得决策成本，推动决策高效决定于执行，因为对于海洋领域来说，政策相对于的影响发挥着巨大的作用，也从另一个方面帮助政府更好地协助企业发展，为沿海企业经济发展提供良好的政治环境与政策引导。政商通力合作，共同推动沿海经济发展。对于沿海渔业来说，每年的沿海休禁渔制度涉及区域广泛，政府通告发布也可以通过该平台进行传播，降低文件沟通成本，合理利用互联网平台进行咨询动态传播，公众也可以通过平台获取最新的海洋相关信息，掌握出海海风、风力等安全信息，保证渔民的财产与人身安全。

如今全球各大主要的海事机构，例如日本海洋数据中心、英国海洋数据中心（BODC）、世界海岛数据库（ISLANDS）、国际海洋数据和信息交换平台（IODE）都在海洋信息收集与共享方面做出了巨大共享，本项目的完成，有助于探索符合我国国情的多源海洋数据的信息平台，通过不同的展示形式，及数据库存储方式，摸索出最合适的信息平台，服务于国家的海洋强国战略。

## 1.3 国内外研究现状

海洋是人类资源最丰富的蕴含区，对海洋资源的有效利用也成为了各国政府的重要战略目标。国内方面：1999年我国正式开始规划“数字海洋”的战略，涉及数字海洋基础性理论研究、数字海洋技术各领域应用和科学工程等要素。2003年国务院批准实施“中国近海海洋综合调查与评价”专项，即“908”专项。2006年中国“数字海洋”信息基础框架构架项目落地执行，中国的“数字海洋”迈入新的台阶，这期间产出了多项实际的应用型项目，其中就包括于2017年正式纳入国家科技基础条件平台中心的国家海洋科学数据共享服务平台，平台由主中心（国家海洋信息中心及分局信息中心）及分中心（中科院海洋所、沿海大学等研究机构）组成，数据共享服务平台旨在充分整合收集国内相关涉海机构丰富的海洋数据信息，基于B/S服务架构，向注册用户提供海洋数据集分类查看，可视化展示，认证下载，接口继承等功能，是一个具备较为完善功能体系的海洋数据共享平台。整体来看，该海洋数据共享平台可以很好地满足公众、企业、研究机构对于海洋数据的需求，可以按照分类进行检索，根据各自的需求进行数据下载，同时对于用户引入了注册登录功能，只有登录并经过机构认证的用户才有资格从网站获取相关数据，保证了数据的安全性。

另一个较为成功的案例为由清华大学牵头开发的海洋大数据分析管理平台，该平台从总体架构上可以分为三个层面，分别是数据层，技术层，应用层。数据层负责接受并处理由数据采集方面（油气平台、水下移动装置、天、空、海岸、船等采集平台）获取的数据，数据内容涵盖遥感数据、地质数据、化学数据等，经过数据层的数据校正、投影变换、数据增强、数据清洗等操作后，进入海洋大数据基础数据平台进行用户层面的展示处理，包括险情预测、数据分析与挖掘、信息集成与检索、海洋信息可视化。该平台的建立有助于集合各科研机构、海事部门、涉海高校、海洋企业分享沟通海量海洋数据，减少数据隔阂，促进海洋数据的重复利用，充分发挥海洋数据在建设海洋强国方面发挥的重要与积极影响。

国外方面：已建立的海洋数据共享平台包括美国国家海洋数据中心、英国国家海洋数据中心（BODC）、国际海洋数据和信息交换平台（IODE）。主要功能包括：公共可获取性数据（涵盖：生物、物理、地理数据），在线申请补充数据，标准化格式及参数化，海洋数据词汇库等。相比较而言，国外的数据共享平台很大程度上优于我们国内现有平台，主要表现在功能更加丰富全面，拥有配套的海洋数据处理软件，对于海洋数据的处理更加完善，不同于国内的数据共享平台只提供按照数据集分类的数据下载。

英国国家海洋数据中心（BODC）可以根据海洋卫星地形图选择一定区域范围内的数据集，同时选择包括划区域选择以及坐标选择（经度、纬度），选定区域或坐标后可以查看当前数据集的数据清洗规则、数据采集及分析方式等内容。对于数据的搜索精度，字段筛选更加丰富详实，用户可以精确到不同区域，不同国家，例如太平洋地区、亚洲地区国家，准确而方便的为用户获取数据提供了便捷。

除去信息共享平台外，海洋数据库的建立也是海洋数据共享平台的前置条件，我们应该秉持数据就是资源的理念，将数据的存储与处理摆在第一位，只有采用合理，具备特殊功能的海洋数据库，才可以将从不同海域，不同设备获取的多样的海洋数据进行有效整合与管理，海洋数据库方面比较成功的案例有全球地形数据集（ETOP），世界海洋数据库（Wrold Ocean Database）

综上所述，国内的海洋数据共享平台与世界先进的海洋数据共享平台在数据数量，丰富程度以及处理方式上仍存在不小差距，值得肯定的是，目前国内几个成熟的海洋数据共享平台已经初具规模，大部分采用的是B/S（浏览器/服务器）模式，降低公众的获取门槛，使得大部分操作可以在云端完成，充分利用了云端处理的高计算能力，降低了本地处理的压力，且保证了数据的上传实时性，降低了因C/S（客户端/服务器）模式而造成的人员成本与更新维护成本，目前看来可以满足公众的需求，和平台最初建立的初衷。

## 1.4 研究内容与章节安排

### 1.4.1 研究内容

本论文的研究主题在与设计并开发符合中国国情与实际的基于多源海洋数据的信息平台，文章对开发平台需要的相关技术理论及浏览器关键技术进行了深入介绍，讨论决定选取何种技术，技术优势所在，并在其后对使用当前技术所面临的开发难点进行了深入讨论，包括使用Flask-SQLALchemy实现数据库对象关系映射，部署服务器至云端，数据文件批量查询分析操作，以期用最合适的技术尽可能优化地实现所设计的功能。其次，本论文结合不同海洋数据库的数据特征，数据处理与分析格式，从关系型数据库和非关系型数据库选择了MySQL数据库作为平台的数据库，既保证了对前台数据展示与搜索的良好支持性，又满足存储海量海洋数据的需求。

本论文以世界上各大海洋机构相关的海洋数据共享平台为考量对象，衡量中国海洋海量数据集内容与数据格式，瞄准国内国际重大海洋数据共享计划，以数据整合共享为目标，以开发为重点工作，最终产出基于多源海洋数据的信息共享平台，服务于社会等研究机构的应用研究包括但不限于海洋渔业、防灾减灾、旅游、航海等方面；协助企业安全高效运营，包括管理海洋风险预警、数据信息化；为政府领导科学决策，海洋管理执法维权提供科学指引。

### 1.4.3章节安排

本论文共分为 个章节，详细内容安排如下所述：

第一章，绪论。简要介绍海洋数据共享平台、海洋数据的有效利用对于发展“数字海洋”的重要意义及影响，该论文的选题及研究意义；最后国内国际海洋数据库发展现状进行了分析和总结。

第二章，系统相关技术研究。

# 2 系统相关技术研究

在进行系统开发前，我们需要详细的了解不同技术的具体应用场景，结合平台需求，选择相应的体系结构和技术。体系结构的差异很大程度上决定着用户使用平台的方式和好感，同时也会对用户体验产生较大的影响。本章将会明确指出我们所使用的Web端开发技术、开发语言、数据库选用等相关内容，包括选取Python作为开发语言的优劣等。

## 2.1Bootstrap前端框架

为了更快更容易的进行Web开发，Mark Otto与Jacob Thornton在Twitter发展出了一套框架，即为Bootstrap，并且将源代码以共享的方式于2011年8月在Github上进行公布。作为一套免费的前端框架，Bootstrap基于HTML和CSS技术设计出了一系列的前端设计模板，包括排版、表格、按钮、导航、模态框、图像轮播和许多其他可选的JavaScript插件，同时引用者也可以在前人的基础上进行自己的设计开发，除此之外，Bootstrap更是拥有强大的响应式设计理念，引用者可以根据实际的应用需求，参考Bootstrap官方文档进行重新布局，以适应不同设备场景下的网页搭建。

使用Bootstrap的好处在于便捷性与低门槛性，任何一个对前端开发技术HTML/CSS以及JavaScript有基本了解，并承担过相应开发工作的程序员，都可以在较短的时间内引入Bootstrap文件进行网页开发；支持响应式布局这一互联网网站最基本的要求，使得Bootstrap迅速成为Github上炙手可热的项目（Bootstrap于2014年6月获得Github项目排名第一的荣誉），这也为我们平台的搭建的稳定性提供了极大的保证；Boostrap作为一款流行前端框架，可以支持所有现有浏览器，包括谷歌浏览器、火狐浏览器、IE浏览器、Edge浏览器、Safari浏览器以及Opera浏览器，可以说能满足我们在用户范围众多的情况下，仍然可以支持不同人群的使用需求；无需下载Boostrap文件至项目目录，可以采用 Bootstrap的CDN（content delivery network：内容分发网络）文件加速功能，通过引入Boostrap的CDN免费加速服务，可以迅速从服务器加载CSS，JavaScript以及图像文件；组件内容丰富，完全可以满足平台的开发要求，且风格相似，有助于实现良好的用户体验。

## 2.2 Web服务器网关接口

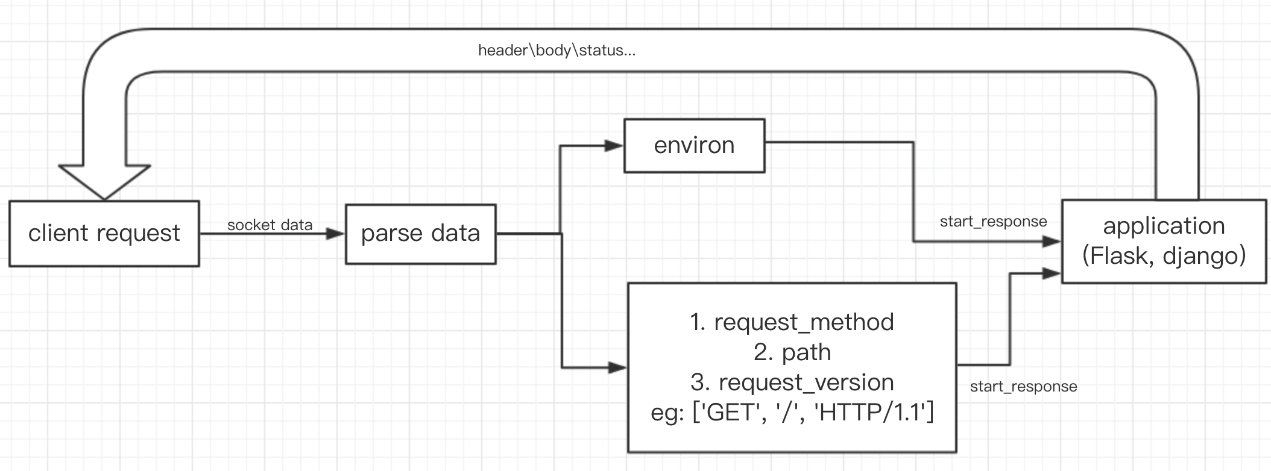
Web服务器网关接口是一类服务于Web服务器与Web应用程序的框架，该接口由Python编写，在这种接口的规定下，Web组件被认为是由client、server、以及middleware组成，而WSGI本身是由服务器（或网关）和应用程序（应用框架）组成，面对一个应用程序传来的WSGI请求时，服务器会根据当前应用程序提供的环境信息产生一个回调函数，一旦应用程序完成了WDGI请求，便可以在回调函数中将数据送入数据库，如此以来可以极大地改善可移植Web应用开发，让开发人员可以集中精力在用Python编写HTML文件，而不需要过于关心HTTP的规范细节，程序员不需要接触到HTTP请求、TCP连接，只需要了解HTTP头部分相关的内容，以及Apache、Nginx、Ligghttpd等静态服务器如何从HTTP请求中读取HTML文件即可，极大地节省了开发人员的时间。

图2-1 WSGI工作原理图

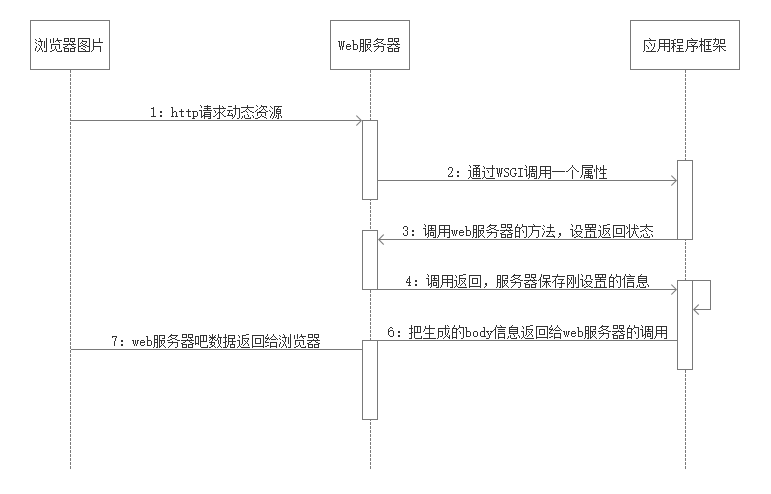


图2-2 浏览器动态请求页面流程图

## 2.3 B/S架构模型

### 2.3.1 B/S架构模型与C/S架构模型的介绍与比较

C/S（Client/Server）即是我们所说的客户机与服务器结构，又称为主从式结构，这一类服务器一般是高性能的PC、工作站、或小型机，客户机是基于内部网 （例如校园、公司、政府单位）且需要安装的专用客户端软件，在这种结构模型中，服务端是一种从属的被动方，它需要得到由客户端传送的请求，然后进行数据处理并传回客户端。而客户端是一种主动的角色，它将需要的操作以请求的方式传送给服务端，当接收到服务端处理后得到的响应后再继续运行。同时，C/S架构一般采用的都是

B/S（Brower/Server）即为浏览器/服务器架构，在这种架构下客户机只需要拥有一个可以上网的浏览器，如谷歌浏览器或者火狐浏览器，服务器安装SQL Serrver或者MySQL数据库，那么就可以使用浏览器通过访问Web Serever访问数据库，进行相关操作。

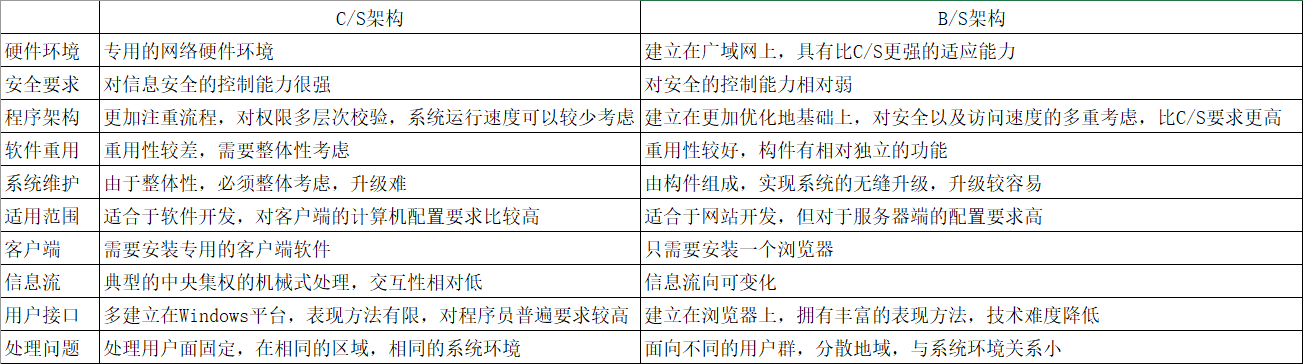


图2-3 B/S与C/S架构对比图

### 2.3.2B/S结构的特点和优势

B/S系统结构基于浏览器进行数据交互，极大的降低了用户获取海洋数据的门槛，用户不需要再想往常一样，必须要到平台的官网下载客户端软件，再通过客户端软件连接数据，从而进行相关操作，此时，用户只需要一个可以连接互联网的浏览器，不用安装任何其他的插件，访问海洋数据共享平台的网址，即可获取实时更新的最新海洋数据，极大地方便了用户的使用 。与此同时，一旦数据共享平台需要发布新的版本，也不需要像B/S架构一样进行客户端版本发布 ，用户手动进行更新，而是开发人员对服务器端进行更新，下一次用户再次访问数据共享平台网址时，就会重新获取最新文件，极大地节省了相关的开发成本。

## 2.4MySQL数据库

### 2.4.1 MySQL数据库介绍

作为目前互联网企业使用量最多，应用范围最广泛的两个大型关系型数据库，Oracle和MySQL都采用了SQL语言进行操作，这一结果决定面对海量海洋数据的高度复杂性以及逻辑性，二者都具有很好的处理能力，本论文选择对MySQL

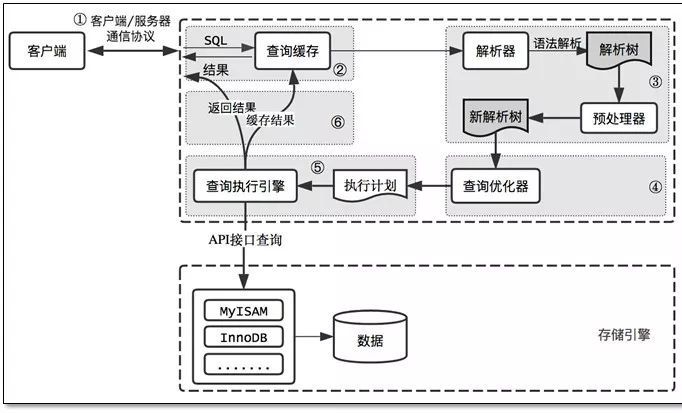
数据库进行分析与讨论，下图展示了MySQL数据库的具体查询过程，可以让我们对MySQL数据库应用于存储各类海洋数据（包括海洋生物、海洋水文、海洋化学等方面）可以起到启发性的作用。

图2-4 MySQL数据库执行流程

### 2.4.2 MySQL数据库所需关键知识

鉴于海洋数据库数据集的复杂关系，不同数据集下拥有不同分类的数据信息，且文件格式不尽相同，这就要求我们可以使用MySQL的主键与外键功能进行数据集与数据文件的关联，以达到快速查询数据文件，以及满足用户可以根据所需要的文件格式进行下载的目的。

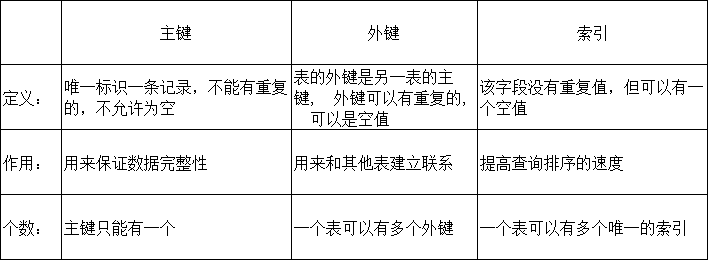
首先，下图给出了MySQL数据库主键、外键以及索引的区别与概念。

图2-5 MySQL数据库主键、外键、索引概念

关系数据库的核心在于SQL语言和二维表的数据结构，所以必须在设计数据库之初就梳理清楚海洋数据集与海洋数据表格的逻辑关系，这样才能避免在前端获取请求进行数据存储、数据处理、数据下载等功能上条理清晰，保证良好的性能。

## 2.5 Python-flask web框架及关键模块

### 2.5.1 Python-flask框架介绍

Python-flask框架是由一个Python的国际化组织Armin Ronacher of Pocoo使用Python语言编写的微量级web框架，之所以被分为微量级框架是因为它的使用不需要额外的引入工具或者库文件，也不需要数据库抽象层以及任何其他的预先存在第三方库的组件。但是，Python-flask却提供了强大的可扩展能力，开发人员可以根据自己的需求添加相应的应用功能以达到增强应用能力的效果，例如对象-关系映射（object-relational mappers）、表格验证（form validation）等功能，目前使用 Python-flask框架的著名企业包括Pinterest（缤趣）、LinkedIn（领英）以及众多采用Python-flask框架构建的门户网站。

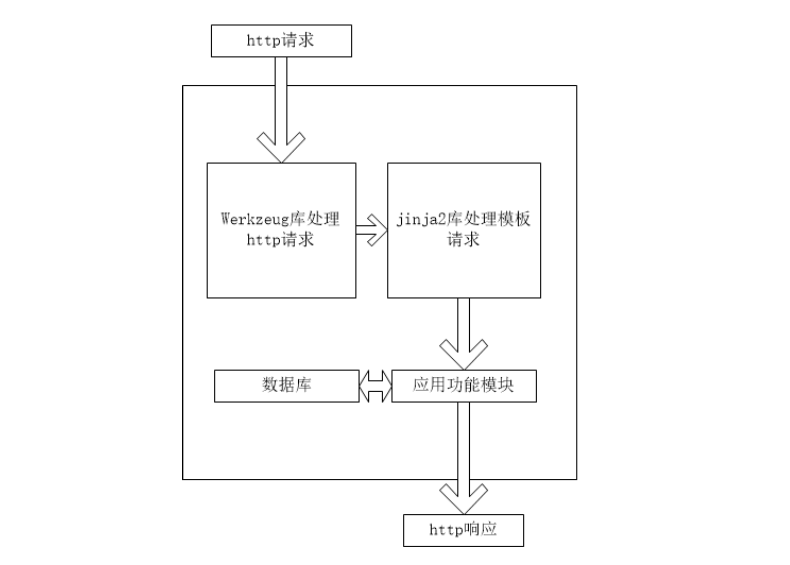
该框架建立于两个Pocoo项目之上：Werkzeug以及Jinja2，Werkzeug是一个专为Python开发语言而编写的实用库，换句话说，也是为WSGI（Web服务器网关接口）应用设计的工具包，采取BSD许可协议，Werkzeug可以实现的软件对象包括请求、回应以及实用函数，开发人员也可以在它的顶层构建一个客户端软件架构，目前支持的Python版本为2.6，2.7以及3.3。Jinja2也是一个采取BSD许可协议的专为Python开发语言编写的web模板引擎，Jinja的作者是奥地利的开源软件开发者Armin，同时他也是我们所采用的flask web框架的原作者。Jinjia2模板引擎与另一个Python web框架Django的模板引擎非常相似，但是Jinjia2同时也提供了类Python的语法表达式，Jinjia2模板引擎允许个性化的定制标签，过滤器、测试和全局变量，不像Django的模板引擎，Jinjia2允许模板的设计者调用带有对象参数的函数，这也为开发人员的功能扩展提供了极大地方便。

图2-6 Python-Flask功能流程图

### 2.5.2 Flask-Uploads模块

Flask-Uploads模块封装的 flask里的文件上传功能，允许应用灵活有效的解决文件上传问题，开发人员可以创建不同的uploads上传集合，每一个集合可以单独的执行上传文件（例如：文件附件上传、图片上传），应用也可以根据不同文件配置相应的URL链接用来存储不同的数据文件，这为海洋数据文件的安全性提供了有效保障。

### 2.5.3Flask-SQLAlchemy扩展

Flask-SQLAlchemy扩展是Python的SQL工具与对象关系映射器，它向应用开发者提供了有力且灵活的SQL控制数据库方法，提供了一整套广为接受的企业级持久性模式，以期可以对数据库进行高效，高性能的访问，并采用简单的Python语言。

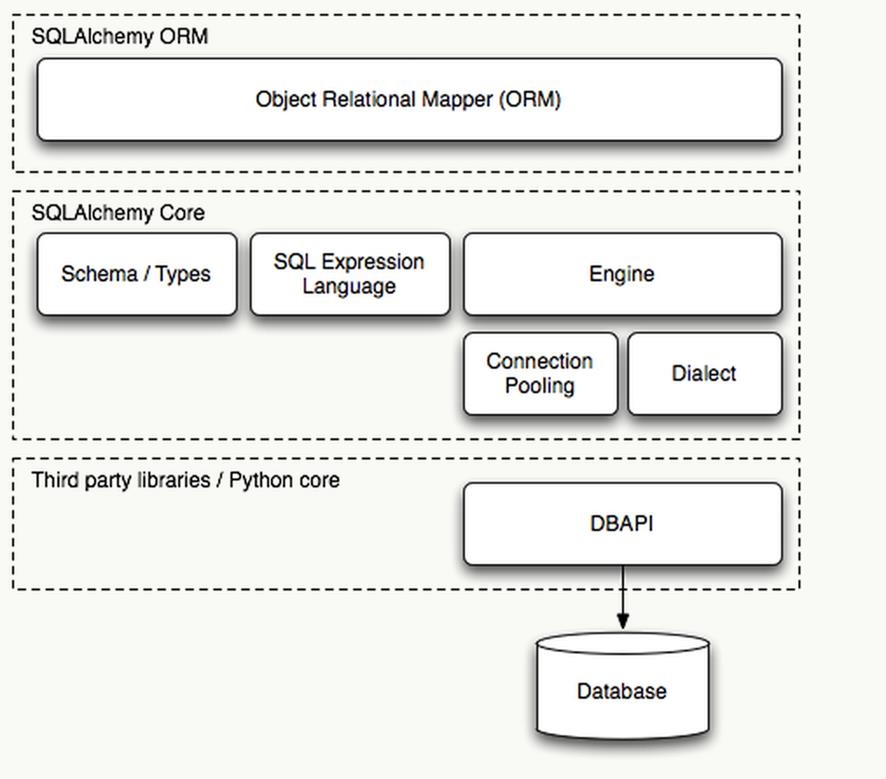


图2-7 flask-SQLAlchemy体系

SQLAlchemy将数据库看作是关系型代数引擎，而不仅仅是一个数据的集合，列不仅可以被选中进行处理，也可以和其他的数据进行联合处理，任何一个小的单元都可以联合成为一个巨大的对象结构，所有的SQLAlchemy都是基于这一核心而展开的。

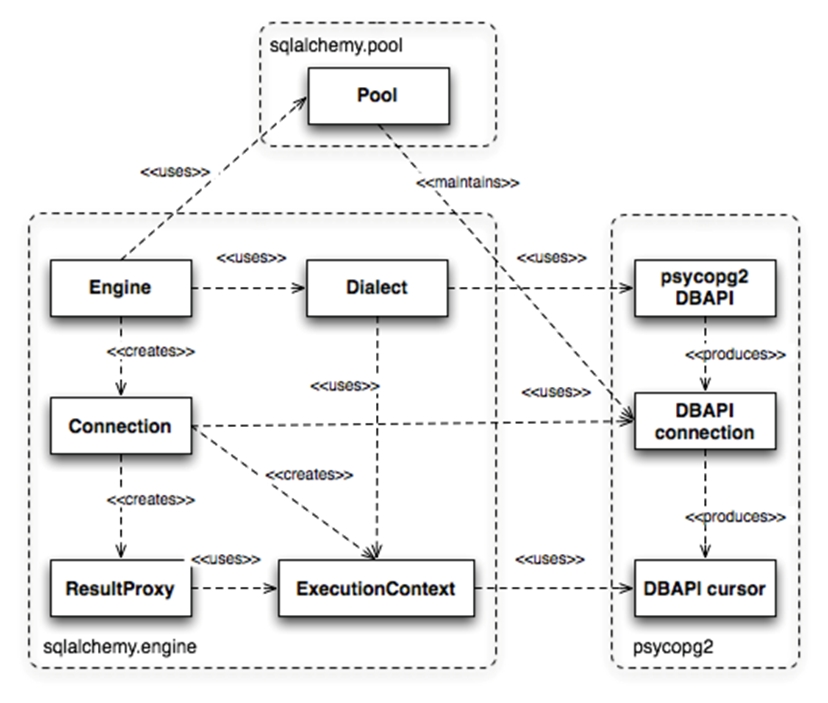


图2-8 引擎连接API

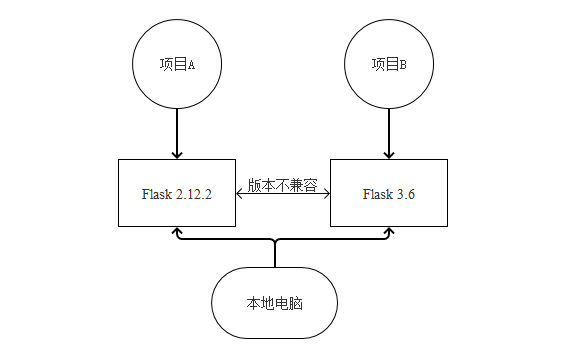
因为Flask-SQLAlchemy是一个扩展程序，所以首先需要激活Python-flask的当前应用环境，之所以要激活当前应用环境是因为Pythond 更新速度非常快，每一个项目的支持版本可能有所不同，所以需要在不同的项目环境中进行激活，避免将来进行操作时产生不兼容的情况。

图2-9 Python-flask虚拟环境示意图

# 3 需求分析及可行性分析

## 3.1系统需求分析

在软件的生命周期中，需求分析阶段的内容是明确平台所需要开发的功能以及非功能需求方面的关键部分 ，对于需要进行开发的功能性需求可借助UML（Unified Modeling Language）中的用例图进行说明，为之后的开发给出详细的性能需求标准。

需求分析有以下几种特点：完整性、准确性、可行性。完整性意味着需求分析应该有始有终，需求分析应该贯穿于软件开发的整个流程，在开发的过程中，设计人员应该实时跟进开发程度，避免因为开发人员与用户的沟通间隙而导致软件与用户最终的目标功能产生较大的偏差。准确性意味着需求分析生成的文档或说明文件应该不会引起开发人员在平台开发过程中产生较大歧义，必须客观正确，既可以清楚明了的描述用户的功能需求也可以减轻开发人员的沟通成本。可行性是指设计人员根据用户需求所确定下的功能性需求应该具有在限定时间内，有限的设备环境下可以完成，而不是异想天开，无法付诸实现。

软件需求的层次：软件的功能与具体性能、用户、业务等组成了软件的需求层次。软件需求需要以软件需求规格说明的形式进行描述，具体的关系见下图所示。

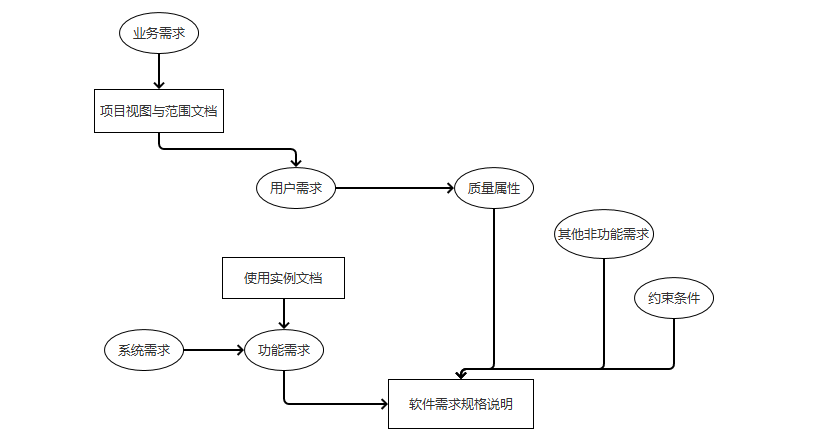


图3-1 软件需求关系

## 3.2平台需求概述

### 3.2.1业务功能需求

平台的开发目的在于收集多源海洋数据，搭建基于B/S架构的海洋数据共享网站，平台的用户包括政府海事部门、涉海高校及研究中心、以海洋为主线进行商业经营的企业，便于政府海事部门依据海量海洋数据做出科学决策，服务军民；为高校及研究中心提供丰富研究资料，降低数据获取成本；为企业提供海洋数据，协助企业高效经营，降低运营成本。

⑴数据查询

主要面向涉海高校与研究中心，海洋数据已经进入大数据时代，数据的更新与收集也成为各个研究机构的首要任务，只有掌握足够多的数据，最新的海洋数据，才能研究发现出解决当下海洋问题（例如：海洋污染、自然灾害）的良策。用户可以在平台根据海洋数据集的关键字进行搜索，内容涵盖：海洋生物、海洋水文、海洋化学等多个方面。

⑵数据下载

主要面向政府海事部门，海事部门对于海洋的管理能力是否达标，在于经略海洋的能力有多大，作为我国自然资源治理的重要组成部分，海洋综合治理能力的提升就在于对海洋数据的掌握程度。在海洋数据共享平台，海事部门可以根据当前决策需要，下载最新的海洋数据，根据数据结合实际分析，定夺最科学的决策方案。

⑶获取资讯动态

虽然互联网发展日新月异，人们获得各类信息的途径也越来越多，但是，鉴于目前还没有形成全国统一的海洋信息发布平台，所以每个地区的信息也无法实现互联互通、互相分享、整合利用，该平台的建设，有力的打破了海洋资讯的“信息孤岛”现象，不但实现政府部门内部的共享，还为沿海地区渔民的出海作业提供了信息平台。

### 3.2.2 平台性能需求

### ⑴响应时间指标

在使用浏览器长时间访问平台的情况下，平台相应指标为：

a.数据查询响应时间≤2秒

b.数据下载响应时间≤2秒

c.页面平均加载时间≤4秒

### ⑵系统容量性能指标

a.支持并发用户数：≤200个

b.支持传输文件大小：≤500MB

c.支持资讯数量：≤1千篇

## 3.3 平台用例分析

### 3.3.1 海洋数据查看与下载

海洋数据模块大体上分为读取与下载两部分，其中读取过程中用户可以根据自己的数据需求进行关键字搜索，具体来说是查看某一个分类下的总数据集时，可以根据数据集名称进行搜索，同时在查看一个单独的数据集时，可以在数据集内置表格中按照数据文件的名称进行搜索。下载业务中，平台要求用户必须要注册账号，使用已注册的账号登录平台，即退出游客模式，进入登录模式，否则只可以浏览查看数据而无法下载任何数据文件。

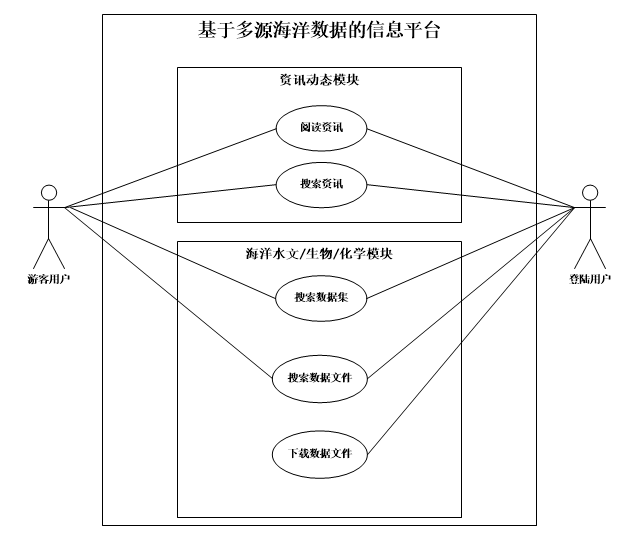


图3-2 海洋数据平台用例图

## 3.4 平台可行性分析

网站采用Python+flask+jinjia2+bootstrap+Mysql，虚拟环境采用Virtualenv解决纯净版本问题。平台借助Bootstrap框架可以为用户提供最好的UI设计，提高用户体验，且采用响应式设计，用户不管在移动端或者说PC端都可以访问网站进行数据浏览与下载。开发平台的可行性研究如下：

### 3.4.1技术可行性

B/S架构相较于C/S架构有更多的灵活性与可扩展性，首先，B/S架构只需要用户安装可以访问互联网的浏览器，键入平台网址即可以快速访问网络，而不需要下载客户端安装包，也免除了日后平台扩展更新的升级操作，一切都在服务端完成，用户无需进行操作。同时，B/S架构的开发方式也更加便捷，开发人员不需要考虑不同用户所使用的终端设备是PC端还是移动端，也不用考虑不同操作系统的差异，所使用的服务器框架flask以及Bootstrap自身都已经对不同浏览器做了适配处理，开发人员只需要专注于平台功能的实现。

Python是一种高级、解释性，通用型动态编程语言，它专注于代码的实现，相较于Java和C++等高级语言来说，Python帮助开发人员用更少的步骤就可以完成功能的实现。在不同的操作系统平台（例如：Unix，Linux，Windows等）上有几乎14%的编程语言选择了Python。为了提高应用的扩展性，使用import语句实现类库的引用，开发人员可以引用flask\_Bootstrap、flask\_uploads等扩展模块。

数据库的选择上，鉴于平台的海洋数据的数据量较大，小型、免费的MySQL数据库成为了我们的首要选择，相比于其他数据库，MySQL不仅可以存储大量的海洋数据，还可以使用SQL这种现代数据库系统通用的语言进行查询操作，同时也为之后使用ORM（Object-Relational Mapping）框架flask-SQLAlchemy进行数据库模型映射提供了必要的前提条件，我们无需像原始的建立数据库的方式那样建立包含多行多列的二维表，而是通过对象的方式，直观的定义海洋数据库。

### 3.4.2经济可行性

经济可行性指的是开发人员在进行平台开发，平台在运行生命周期内的维护支出（包括服务器费用支出，域名费用支出、数据库费用支出等），同时我们需要考虑用户在使用平台的过程中所产生的成本，既包括时间成本，也包括经济成本。经过分析，开发成本方面，因为该项目属于校内学术性项目，所以可以免费试用MySQL数据库，所使用的也是开源的python-flask框架，开发人员可独立完成平台的建设工作。用户成本方面，用户访问网站时，只需要花费少许移动流量，且在大文件数据下载时，可以从表格中看到当前所下载大文件数据的文件大小，以便用户衡量是否选择消耗流量进行下载。

# 4 平台总体设计

## 4.1 平台架构设计

从应用角度来看，系统架构主要表现在用户可以通过访问平台的网站链接来进行数据的查看，数据的关键词搜索以及下载操作，从软件的逻辑角度来看，整个软件平台结构如图所示，连线在此处指的是调用关系，即下层模块对上层模块的支持。

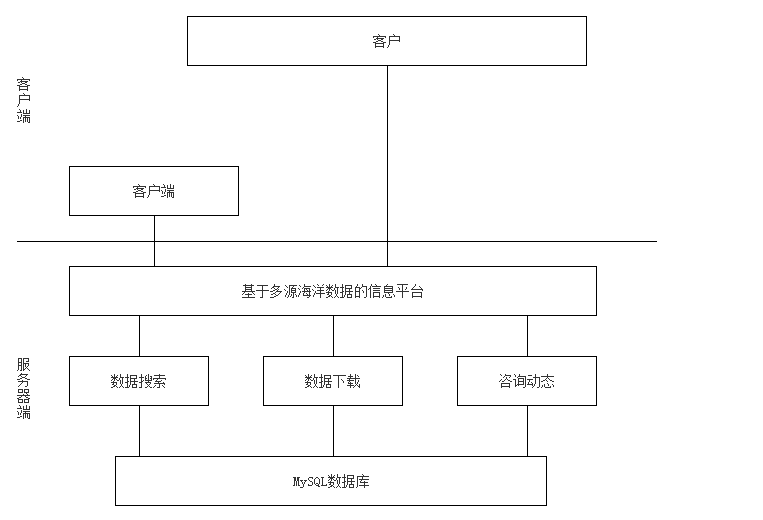


图4-2 海洋数据信息平台软件架构

图3-3描述了海洋数据信息平台的功能架构，平台可运行在Windows和Linux环境中，由Python-flask框架提供Web服务，使用MySQL数据库，以网站的形式为用户提供服务。

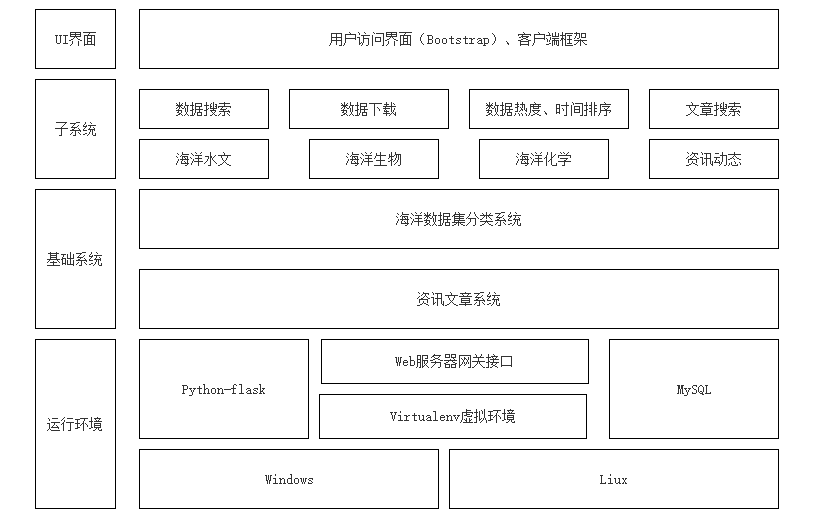


图4-3 海洋数据信息平台功能架构

## 4.2平台的框架设计

海洋数据平台采用WEB开发中最常使用的ORM框架，ORM框架是对象-关系映射（Object-Relational Mapping）的简称，目前各大互联网公司中应用开发环境的主流开发方法就是面向对象的开发方法，而关系数据库是数据处理与存储的主流存储系统。而在Python语言开发过程中，使用范围最广的ORM框架就是SQLAlchemy，使用SQLAlchemy可以相当程度上减少SQL语言操作MySQL数据库的繁琐过程，而是采用更为形象的对象模型来映射数据库，既可以达到操控数据库的目的，又可以清楚明了的通过调用对象模型的方式进行增删改查等数据库常规操作。

## 4.3 平台的模块设计

对平台架构和框架设计完成后，下一步的工作就是根据具体的需求，结合现今的海洋数据现状，对平台的功能模块进行设计，这一步决定了用户的需求在平台的具体展现方式，以及用户可以使用的所有功能。

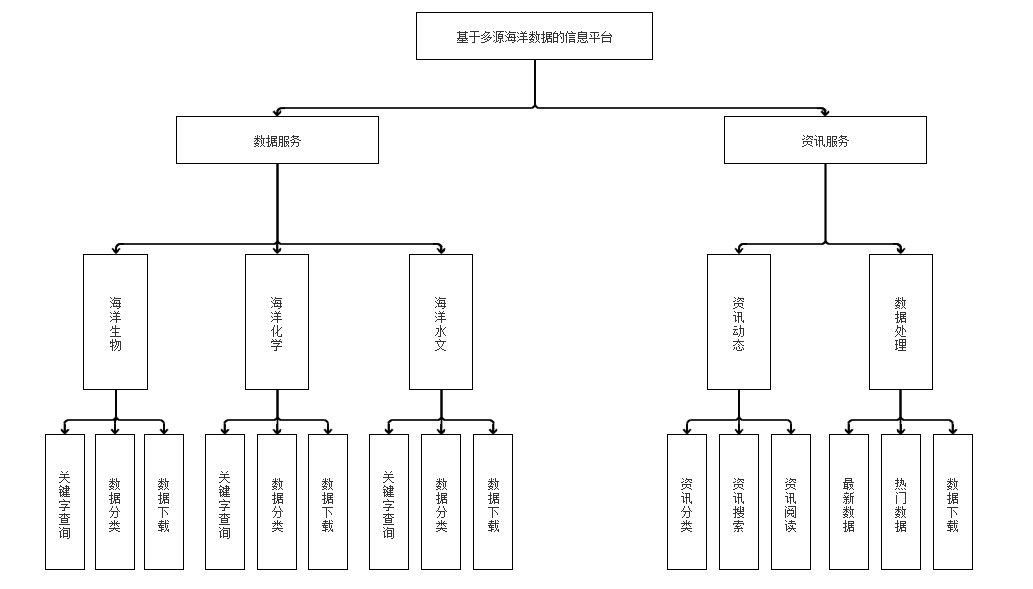


图4-4 海洋数据信息平台模块设计

从图4-4平台模块设计图中可以看出来，具体的海洋数据信息平台模块设计包括海洋生物、海洋化学、海洋水文、资讯动态、数据处理。

1.海洋生物数据集功能：主要功能有海洋生物数据的分类展示，根据用户的数据需求，用户可以根据数据集名称进行关键字搜索；数据按照数据集的分类进行收集，以表格的形式展示属于同一数据集的数据，展示内容包括数据编号、数据名称、上传时间、数据格式。用户在数据表格中可以再次根据需要的数据按照数据文件名称进行搜索，同时可以进行下载操作，文件流作为一个输出的流文件，放在response中返回给浏览器，浏览器从response中取出session内容，进行文件下载。

2.海洋水文数据集功能：主要功能有海洋水文数据的分类展示，依据数据集关键字进行模糊搜索、数据下载，每一个数据集展示表格的字段把包括数据编号、数据名称、上传时间、数据格式，同时用户也可以根据自己的数据需求进行数据下载。

3.海洋化学数据集功能：主要功能有海洋化学数据的分类展示，不同数据集的卡片式展示、展示的字段包括数据集大小、数据集来源。每一个数据集用户都可以根据需求进行数据文件模糊搜索（根据文件名称），下载需要的数据文件。

4.资讯动态功能：主要功能包括资讯动态的分类，卡片展示，查看文章，文章的字段包括发布来源、文章类型、发布时间。同时提供用户根据文章标题进行模糊搜索。

## 4.4 平台数据库设计

对于存储多源海洋数据的数据平台而言，整个平台功能的实现都依赖于在数据库中存储的海洋数据信息，包括海洋生物数据、海洋化学数据、海洋水文数据、资讯文章数据，依据标准的数据库设计工作流程，在进行数据库的具体设计之前，需要先借助系统设计中经常使用的ER模型，明确在数据库中存储的海洋数据实体信息，然后根据实际用户的需求进一步分析每一个实体应该具有的属性，掌握清楚以上内容后，再根据数据库实体信息和具有的属性进行海洋信息数据库的表格设计，给出每一个实体在对应数据库中存储的数据表格信息。

### 4.4.1 实体描述

在对数据库实体进行分析后，我们可以得出海洋数据信息平台总体的E-R图，实体、属性和联系是我们在设计过程中必须要考虑三个要素，其中实体指的是首要的数据对象；属性是作为描述信息，用来详细的说明一个数据对象；关系表示实体之间的联系，其中关联关系又分为三种方式：一对一、一对多、多对多，海洋数据平台的E-R图详细内容如图所示。

对海洋数据平台进行数据库设计的时候，最重要的任务就是确定海洋数据平台在数据库实体方面的内容，主要依据就是目前所掌握的海洋数据文件格式、海洋数据类别、不同用户对于数据的需求大小、整个平台的业务流程等等。进而明确每一个数据库实体及其所拥有的属性，不同实体之间的关系。经过E-R图的分，我们可以明确数据库实体有海洋化学数据、海洋生物数据、海洋化学实体、资讯动态实体、用户实体等内容。

### 4.4.2 数据库迁移

数据库设计也要与开发进度进行协调，因为数据库的设计不可能一开始就全部确定下来，所以我们需要动态的进行数据库调整，将不合理的数据库设计进行更改，而采用SQL工作量较大且容易发生错误，一旦数据库发生错误，就极容易对数据库中存储的数据造成不可恢复的影响，在真正的生产环境中，总会发生数据库修改的行为，但是，一把我们不会直接手动去修改，而是采用数据库迁移的方式修改ORM对应的数据库模型，然后采用工具进行数据库迁移、数据库表更新，本项目所采用的工具就是flask-migrate，flask-migrate是基于Alembic进行的一次封装，并可以通过Flask以扩展的方式进行使用，但最根本数据库迁移操作其实都是Alembic进行的，它可根据所对应的数据库模型的变化而变化，并及时将这种变化反应到MySQL数据库中。

### 4.4.3 数据库表格设计

数据库的设计不仅要参照用户的需求进行考虑，衡量我们现今所拥有的海洋数据，具体包括海洋生物数据、海洋水文数据、海洋化学数据，以及每个海洋数据分类集下的不同格式文件，最后还应当考虑在平台的业务体系下，数据库应当以最好的性能服务用户。所以我们要采用科学的设计方案和流程进行数据库的设计。

基于多源海洋数据的信息平台包括大功能模块，因此我们创建了个数据库模型，分别映射到数据库的不同功能模块，以下分别对不同的功能模块进行介绍。

a.用户信息表

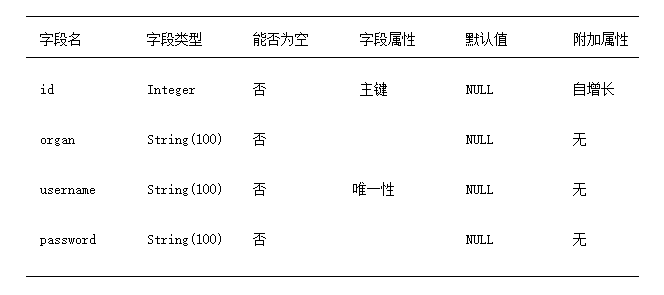
用户信息表主要用来保存在平台通过注册的用户，保存的用户字段包括用户主键ID，用户所属单位，用户名，密码。游客模式下，用户可以查看平台的海洋水文、海洋生物、海洋化学数据，但只有登录模式下的用户，才拥有下载文件的权利。同时主键ID作为在user表中唯一标示用户的字段，可以用来和其他表建立外键链接。用户信息表详情见下图：

图4-5 用户信息表

b.轮播图表

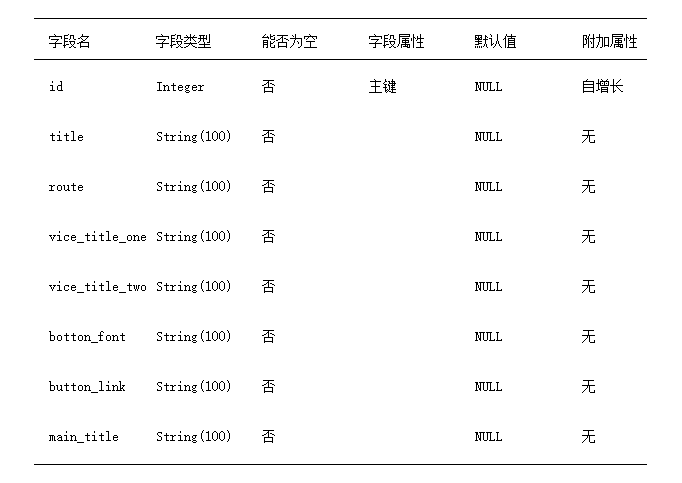
轮播图表主要用来存放平台首页index.html动态轮播图的显示内容，为了方便平台将重要的信息展示与此，并可以及时进行更新，而不需要进行二次开发，提高了程序的灵活性。保存的轮播图表字段包括轮播图主键ID，轮播图图片描述，轮播图图片路径（相对于文件夹的路径），轮播图大标题，轮播图副标题一、副标题二、轮播图按钮文字、轮播图点击跳转URL地址。轮播图表详情见下图：

图4-6 轮播图表

c.资讯动态表

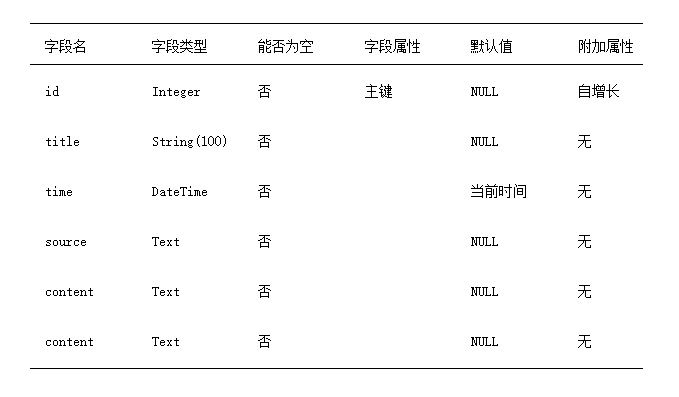
资讯动态表主要用来平台的所有资讯文章，游客模式用户和已登录用户均可以查看相关动态，及时了解海洋相关资讯。保存的资讯动态表字段包括资讯标题、资讯类型、资讯发布时间、资讯发布来源、资讯正文内容（HTML文本），资讯动态表详情见下图：

图4-7 资讯动态表

d.海洋生物表

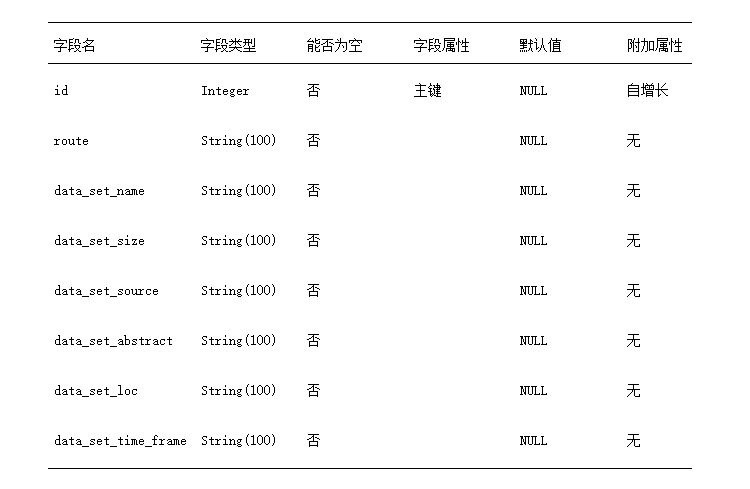
海洋生物表主要用来存放海洋生物数据集的相关信息，用户如果要寻找某一类数据，可以先从大类出发，例如从海洋生物、海洋化学、海洋水文数据集寻找，判断自己要寻找的数据属于哪一类数据，然后再根据大类数据寻找数据文件所属数据集，海洋生物表所存储的字段包括海洋生物数据集主键ID、数据集图片路径、数据集名称、数据集文件总大小、数据集来源、数据集概述、数据集覆盖范围、数据集覆盖时间段。海洋生物表详情见下图：

图4-8 海洋生物表

e. 海洋生物数据文件表

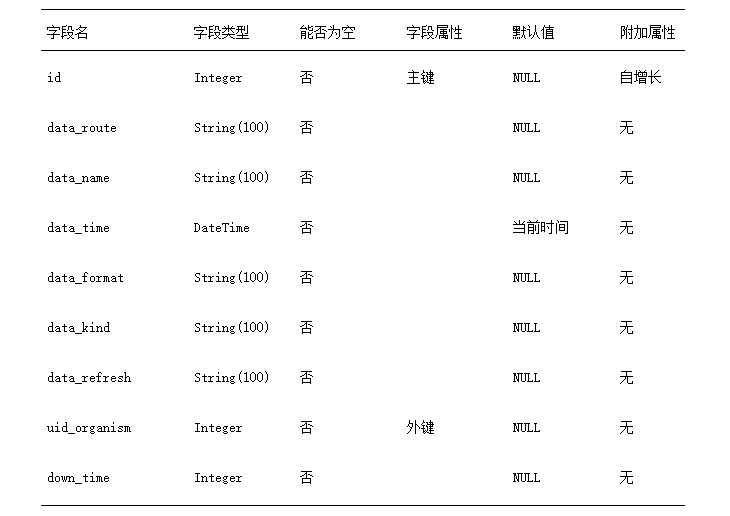
海洋生物数据文件表主要用来存放一个海洋生物数据集下面所属的文件，将数据文件与数据集分开建表，不仅有利于平台相关人员对数据的管理，有加快了数据的查找速度，提高了用户下载文件速度，有效地改善了用户体验。海洋生物数据文件表包括数据文件主键ID 、数据文件在项目中的存放路径（data\_route）、数据文件的上传时间、数据文件后缀格式、数据文件所属数据集、数据文件更新频率、数据文件所对应的数据集的外键、数据的下载次数。

图4-9 海洋生物数据文件表

除了海洋生物表和海洋生物数据文件表之外，还有两套同样字段的表格，即海洋水文表、海洋水文数据文件表和海洋化学表、海洋化学数据文件表。因为字段相同，此处不再赘述。

# 5 基于多源海洋数据信息平台的实现

本章根据前文所述需求分析及可行性分析，在平台总体设计的指导下，着手创建Python项目，使用Python-flask开发框架、Navicat Premium（可多重连线数据库的管理工具）、PyCharm（用于Python语言开发的计算机编程集成开发环境）等。完成了基于多元海洋数据的信息平台的实现。

## 5.1 项目前期工作

### 5.1.1安装组件

因为平台开发使用的是Python-flask框架，为了提高开发效率和平台的性能，我们选择了Python-flask中提供的众多Web开发组件与对应的扩展库，以下列出开发本平台时候所使用的组件，扩展库及其功能。



图5-1 平台所使用组件及扩展库

### 5.1.2 数据库创建与迁移

Flask-migrate扩展提供了通过修改ORM来控制数据库的功能，因为创建数据库后，在后期修改字段的时候，不会自动的映射到数据库中，必须删除表，然后重新运行相关命令才会重新映射，这样不符合我们的需求。因此flask-migrate就是为了解决这个问题，它可以在每次修改模型后，都将修改的东西映射到数据库中。

flask-migrate相关的命令如下：

python manage.py db init：初始化一个迁移脚本的环境，在最开始建立环境时运行，且只需要执行一次。

python manage.py db migrate：将模型生成迁移文件，此后只要模型有所改变，都要运行当前命令进行对象映射，以同步数据库为最新状态

python manage.py db upgrade：将迁移文件真正的映射到数据库中。每次运行了migrate命令后，都要紧跟着运行这个命令。

## 5.2 系统主要功能模块的实现

### 5.2.1 平台登录模块的实现

使用该平台的用户不必安装本地的客户端软件，只需要使用可以正常连接上网的浏览器，输入平台的URL地址，即可以跳转至平台的登录页面，在登录页面，用户需要输入自己的账号和密码进行验证登录，服务器将拿到的账号密码与数据库中存放的用户信息的user表进行比对，首先寻找用户名相同的用户，如果没有找到则代表该用户没有注册，无法登录，此时用户可以点登录框下方的“点击注册”按钮进行在线注册；如果在user表中找到与用户所输入的用户名相匹配的用户信息，则开始匹配该用户在user表存储的密码是否与用户所输入的密码相符合，如果符合，则代表用户信息验证成功，浏览器页面跳转至平台首页，用户可以进行浏览数据，下载数据等操作。系统登录功能关键代码如下：



具体的登录界面效果图如图所示：

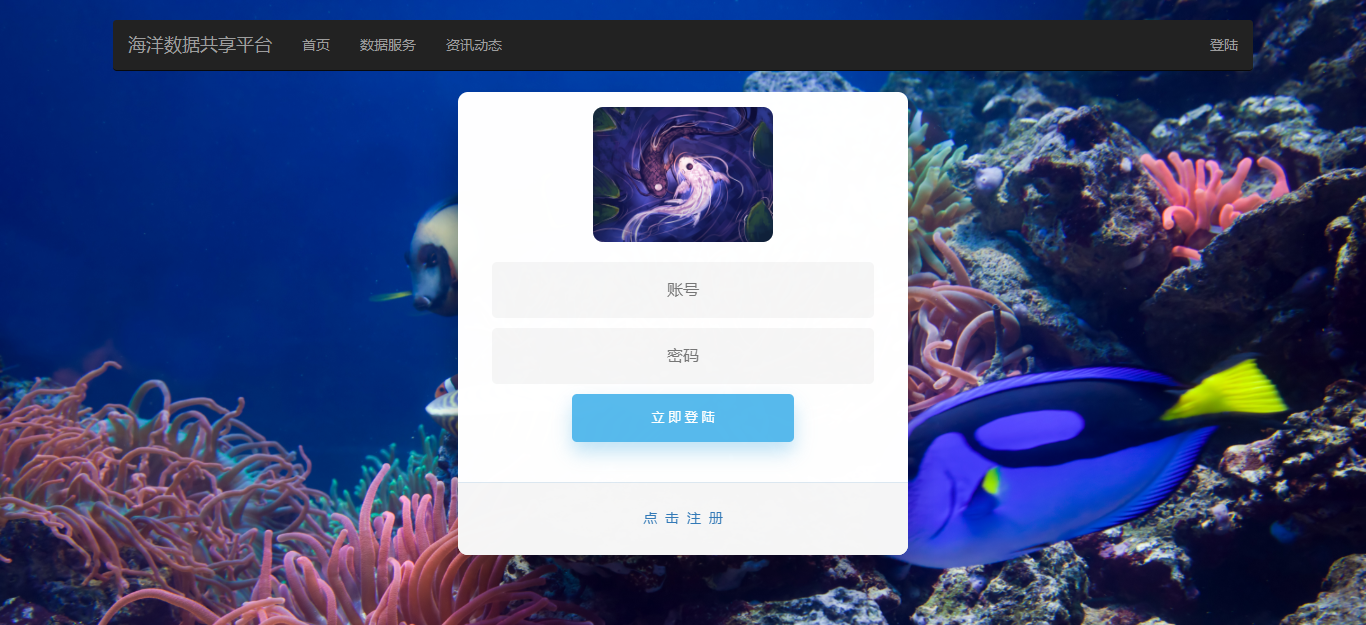


图5-2 平台登录界面效果图

### 5.2.2 平台注册模块的实现

平台注册模块主要用于实现游客用户进行平台账号注册，只有在此处进行注册，并成功将信息存储在user表中的用户，才可以登录进行海洋数据文件下载。在用户注册的过程中，平台设置了相关的注册要求，以确保用户账号密码的安全性，具体设置要求如下：

用户名：

用户名长度不能少于5个字符

用户名只能为英文数字和下划线，且不能以数字作为开头

密码：

密码长度不能少于8个字符

密码中至少包含字母、数字、特殊字符的两种

两次输入的密码必须一致

验证码：

验证码必须正确填写（不区分大小写）

所属单位：

所属单位不能为空

只有满足上述所有要求的情况下，用户才可以点击注册按钮进行注册，否则相应的输入框会给出提示信息，直到用户所输入的信息满足校验要求。所对应的不同样式如下图：



图5-3 注册信息填写（正确）效果图



图5-4 注册信息填写（错误）效果图

当用户提交了所有正确信息后，平台会将获取到的账号与user表的username字段进行比对，判断是否数据库已经存在该用户，如果匹配成功，则代表用户已经在平台进行了注册，此时置空所有字段，并且显示提示信息，提供登录链接给用户，用户可以直接点击登录按钮进行跳转，在登录页面进行再次登录。

图5-5 用户名已注册效果图