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

# 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

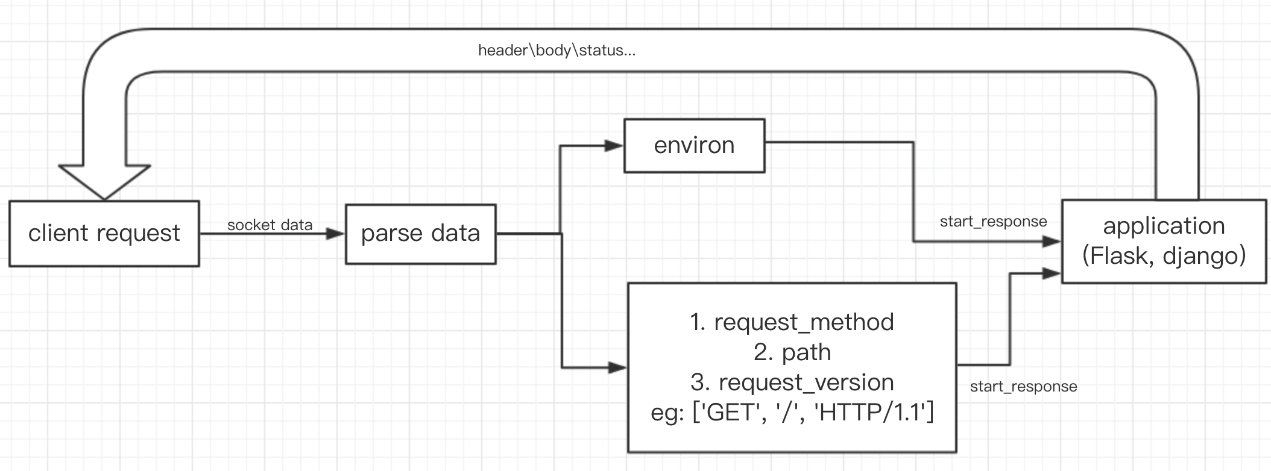
Web服务器网关接口是一类服务于Web服务器与Web应用程序的框架，该接口由Python编写，在这种接口的规定下，Web组件被认为是由client、server、以及middleware组成，而WSGI本身是由服务器（或网关）和应用程序（应用框架）组成，面对一个应用程序传来的WSGI请求时，服务器会根据当前应用程序提供的环境信息产生一个回调函数，一旦应用程序完成了WDGI请求，便可以在回调函数中将数据送入数据库，如此以来可以极大地改善可移植Web应用开发，让开发人员可以集中精力在用Python编写HTML文件，而不需要过于关心HTTP的规范细节，程序员不需要接触到HTTP请求、TCP连接，只需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

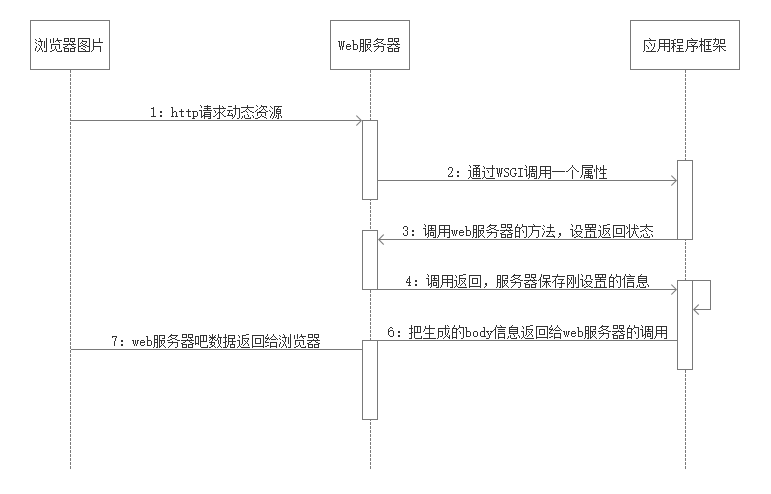


图2-2 浏览器动态请求页面流程图

## 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。

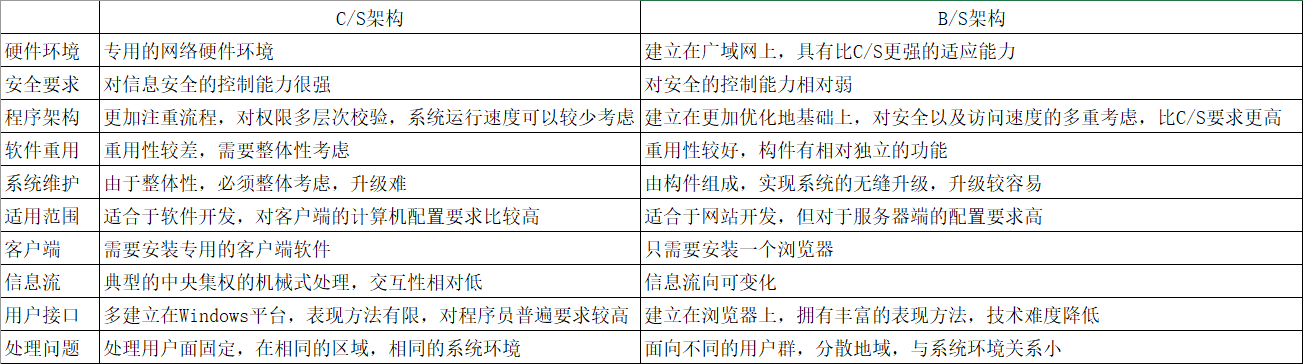


图2-3 B/S与C/S架构对比图

### 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

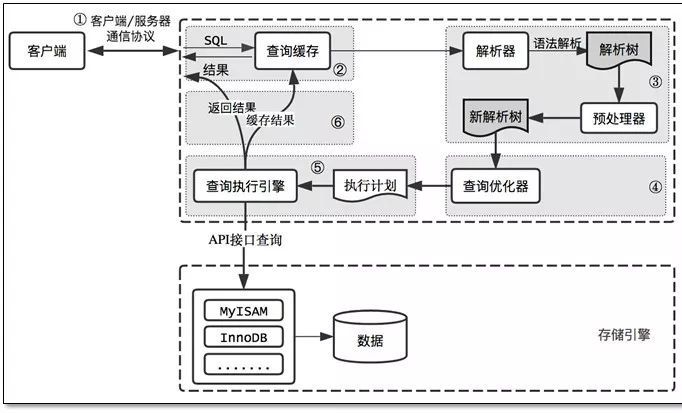
数据库进行分析与讨论，下图展示了MySQL数据库的具体查询过程，可以让我们对MySQL数据库应用于存储各类海洋数据（包括海洋生物、海洋水文、海洋化学等方面）可以起到启发性的作用。

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。

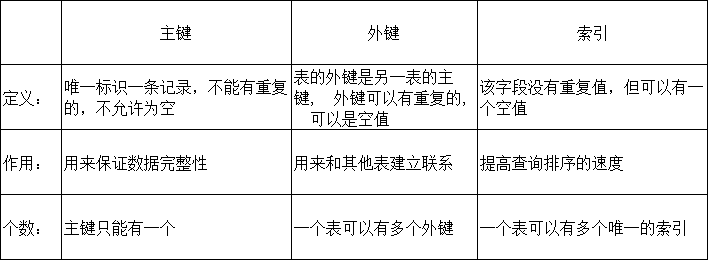
首先，下图给出了MySQL数据库主键、外键以及索引的区别与概念。

图2-5MySQL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

### 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。