

通用生态观测数据管理平台的设计与实现

陈灿

chencan@cnic.cn

科研信息化实验室

中科院计算机网络信息中心

背景

第一阶段

采集方式

野外通过纸笔记录

第二阶段

野外通过订制的笔记电脑或PDA手持终端

数据存储和管理方式

录入到电脑，以word、excel等文件为载体

导出成word、excel等文件或导入到关系型数据库

第三阶段

Android等系统的手持终端

直接上传关系型数据库，并有数据管理系统辅助

现状

- ❖ 采集方式和数据管理平台
 - ❖ 专业设备加笔记本
 - ❖ 基于PDA，配合关系型数据库
 - ❖ 基于Android平台，针对特定业务领域，配合关系型数据库
- ❖ 异构数据处理方法
 - ❖ 已有异构数据库基础上，通过XML等中间件进行集成查询
 - ❖ 采用各种中间件，整合原有异构数据，基于大数据技术，重新构建新数据库

问题

- ❖ 跨学科异构数据的融合存储
- ❖ 高度可扩展的数据存储模型
- ❖ 支持I/O密集操作的数据管理平台

```
        "desc" : "潮汐情况",
        "attribute" : "string",
        "empty" : "0",
        "value" : ""
    },
    {
        "name" : "remark",
        "desc" : "备注",
        "attribute" : "string",
        "empty" : "0",
        "value" : ""
    },
    {
        "name" : "total_species_num",
        "desc" : "总种数",
        "attribute" : "double",
        "empty" : "1",
        "value" : ""
    },
    {
        "name" : "total_num",
        "desc" : "个体总数",
        "attribute" : "double",
        "empty" : "1",
        "value" : ""
    }
],
"detail" : [
    {
        "name" : "species_name_C",
        "desc" : "中文名",
        "attribute" : "string",
        "empty" : "1",
        "value" : ""
    },
    {
        "name" : "species_name_S",
        "desc" : "学名",
        "attribute" : "string",
        "empty" : "1",
        "value" : ""
    },
    {
        "name" : "adult_num",
        "desc" : "成体数",
        "attribute" : "double",
        "empty" : "1",
        "value" : ""
    },
    {
        "name" : "young_num",
        "desc" : "幼体数",
        "attribute" : "double",
        "empty" : "1",
        "value" : ""
    }
]
}
```

类目	
观测单位	
观测者	
观测日期	
观测样地名称	
观测样地代码	
观测样地类型	森林 (
地理位置	经度: _____ 纬度: _____
观测样地建立时间	
样地面积和形状	
气候条件	
地形地貌	海拔
	地貌状况
	坡度
	坡向
	坡位
土壤状况	土壤类型
	土壤母质
	土壤剖面特征
植被状况	区域植被类型
	群落类型
	群落层次结构及各层优势物种
	演替阶段
动物活动状况	
人为干扰活动类型和强度	

填写说明：

1) 地貌状况描述参见附录 K。

2) 人为干扰活动类型和强度参见附录 0。

- ❖ 采集数据模型
- ❖ 两类采集表数据
- ❖ 同一数据格式
- ❖ BSON文件 key

```
        "_id" : ObjectId("5791d290b56e605f0ea5354f"),
        "org_name" : "系统测试",
        "table_name" : "t_nl_fqzsfjlb",
        "person" : "付海静",
        "p_group" : "hahaha",
        "born_time" : "20160719_200000",
        "upload_time" : "2016/7/22 16:0:16",
        "brief" : {
            "head" : {
                "date" : "日期",
                "temperature" : "温度",
                "recorder" : "记录者",
                "altitude" : "海拔",
                "longitude" : "经度",
                "latitude" : "纬度",
                "time_start" : "开始时间",
                "time_end" : "结束时间",
                "habitat_type" : "生境类型",
                "human_ac_type" : "人为干扰活动类型",
                "human_ac_strength" : "人为干扰活动强度",
                "remark" : "备注",
                "total_species_num" : "总种数",
                "total_num" : "个体总数"
            },
            "content" : {
                "date" : "2016-07-19",
                "temperature" : "23",
                "recorder" : "付海静",
                "altitude" : "2222.22222",
                "longitude" : "18.23542",
                "latitude" : "168.12542",
                "time_start" : "2016-07-19 15:23:23",
                "time_end" : "2016-07-19 18:23:23",
                "habitat_type" : "针阔叶混交林",
                "human_ac_type" : "狩猎",
                "human_ac_strength" : "中",
                "remark" : "备注",
                "total_species_num" : "12",
                "total_num" : "434"
            }
        },
        "detail" : {
            "head" : {
                "species_name_C" : "中文名",
                "species_name_S" : "学名",
                "adult_num" : "成体数",
                "young_num" : "幼体数"
            },
            "content" : [
                {
                    "species_name_C" : "某种鸟",
                    "species_name_S" : "asdfawegdr",
                    "adult_num" : "34",
                    "young_num" : "12"
                }
            ]
        }
    }
```

通用生态监测数据管理平台

用户交互层：

采集表定制系统

采集数据管理系统

采集轨迹管理系统

.....

功能模块层：

文件操作

数据库操作

Session处理

页面路由控制

权限管理模块

.....

驱动层

ExpressJS

NodeJS

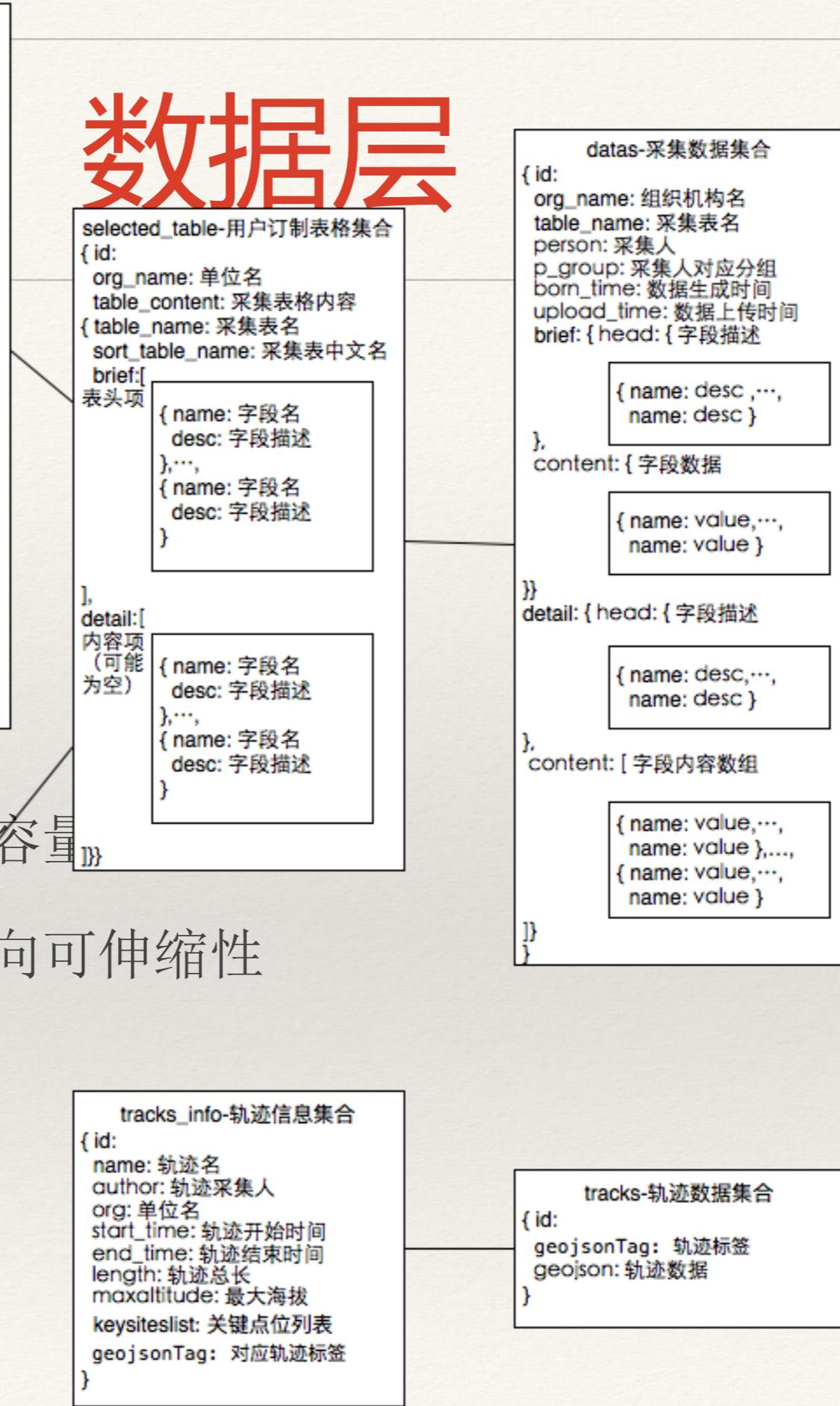
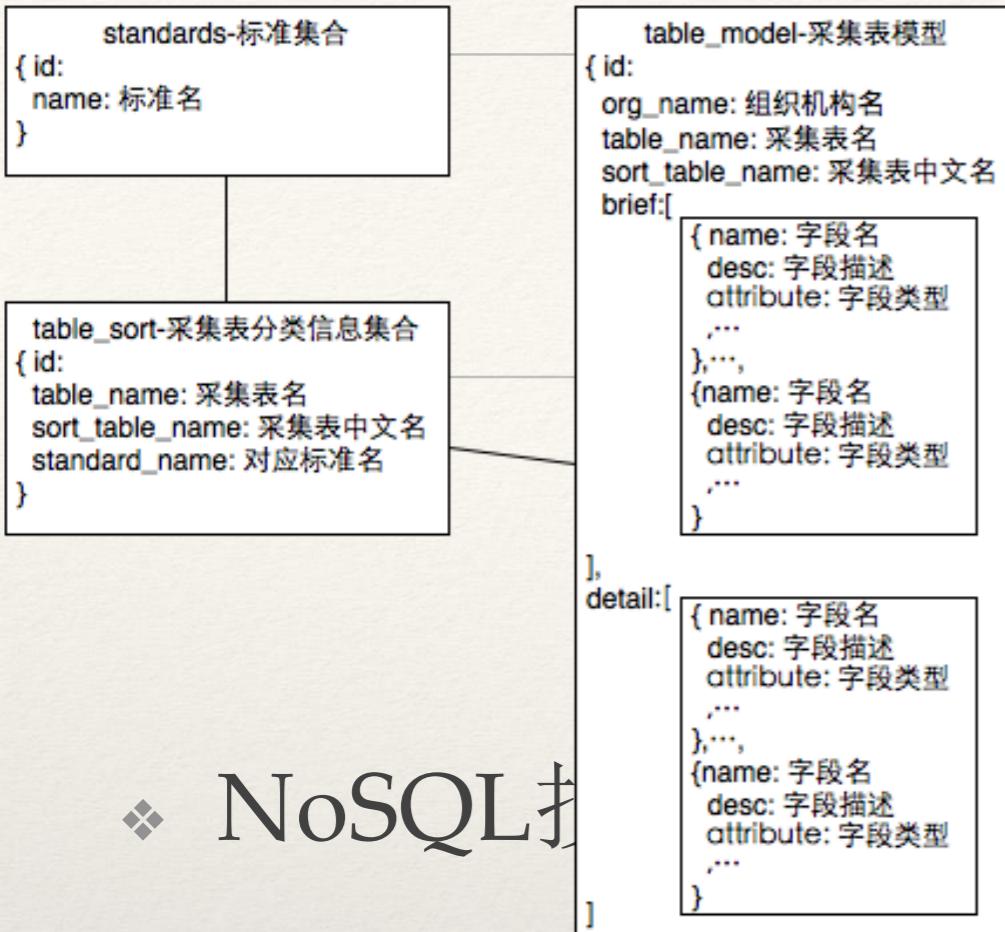
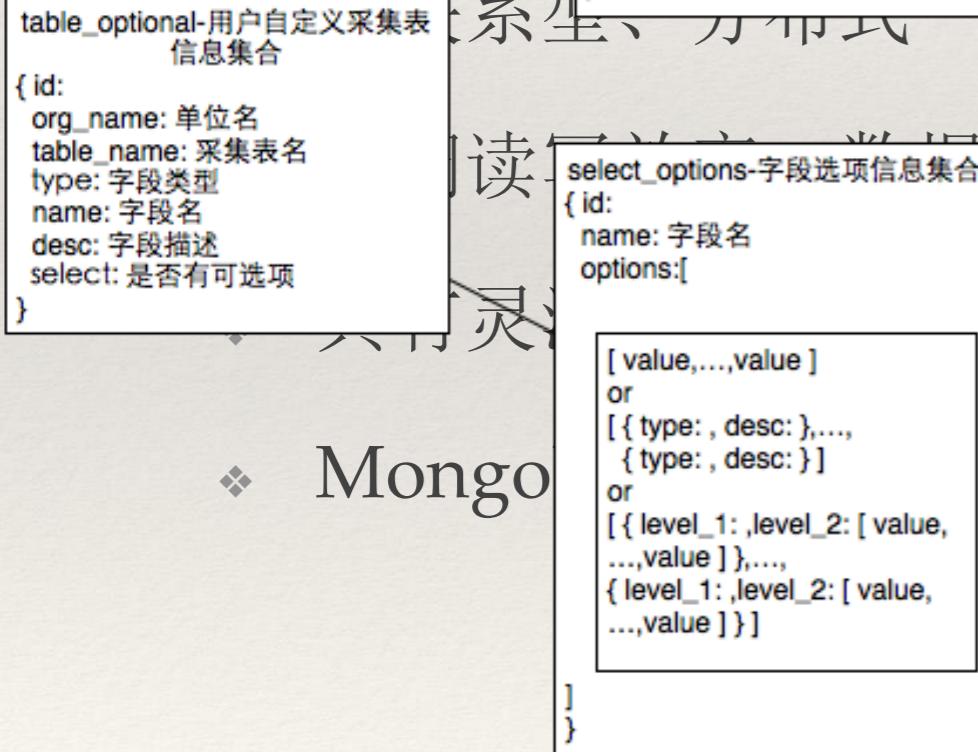
数据层：MongoDB

JSON格式采集表模板

XML格式采集记录

移动终端数据采集系统

◆ NoSQL



驱动层

- ❖ 平台特点
 - ❖ 多用户使用
 - ❖ 针对大数据存取
 - ❖ I/O密集
 - ❖ 高并发
- ❖ Node.js
 - ❖ 后端的Javascript运行环境， Google Chrome的V8引擎， 提供系统级的API
 - ❖ 事件驱动、非阻塞I/O模型

功能模块层

- ❖ 标准化模块
- ❖ 文件操作模块
- ❖ 数据库操作模块
- ❖ Session处理模块
- ❖ 页面路由控制模块

用户交互层

- ❖ 采集表订制管理系统
- ❖ 数据管理系统
- ❖ 轨迹数据管理系统
 - ❖ 生态观测采集表的订制
 - ❖ 采集数据上传
 - ❖ 数据入库前审核
 - ❖ 数据管理
 - ❖ 轨迹数据上传和展示

移动采集端-平台

- ❖ 软件平台-Android
 - ❖ 兼容性
 - ❖ 丰富的功能选择
 - ❖ 多进程
 - ❖ 运营商支持 设备种类多 硬件配置高

移动采集端-功能

- ❖ 数据采集
- ❖ 轨迹记录
- ❖ 账户管理
- ❖ 系统设置

移动采集端-截图

野外数据采集系统

当前：内蒙古自治区(21) 内蒙古自治区

已下载：-

热门城市

- 北京市(131) -24.4M
- 上海市(289) -30.1M
- 广州市(257) -21.9M
- 南京市(315) -10.6M
- 杭州市(179) -13.5M

全国

- 全国概略地图包(1) -27.8M
- 北京市(131) -24.3M
- 上海市(289) -30.1M
- 天津市(332) -14.8M
- 重庆市(132) -34.9M
- 安徽省(5) -55.6M
- 福建省(6) -40.2M
- 甘肃省(7) -40.3M
- 广东省(8) -126.4M

轨迹查询

以下请输入一项或多项(全空查询全部轨迹)

选择分组或人员:

相关的地点:

时间段查询:

开始时间 到 结束时间

区域轨迹查询

选择矩形区域方法: Windows系统--Ctrl+拖拽鼠标; Mac OSX系统--Command+拖拽鼠标

轨迹列表

轨迹总长度: 208114米 显示所选

ID	名称	开始时间	文件大小	操作
<input checked="" type="checkbox"/>	1 居民登记	2015年8月11日 GMT+8 上午 12:15:02	18.63 MB	
<input checked="" type="checkbox"/>	2 小巡护	2015年8月10日 GMT+8 下午 7:06:21	2.41 MB	
<input type="checkbox"/>	3 居民登记 (阿尔山保利格)	2015年8月6日 GMT+8 上午 4:13:41	15.22 MB	
<input type="checkbox"/>	4 居民登记 (阿尔山保利格)	2015年8月6日 GMT+8 上午 6:11:15	40.27 MB	
<input type="checkbox"/>	5 西旗采购和加油	2015年8月4日 GMT+8 下午 11:30:14	3.83 MB	

显示第 1 到 第 5 条记录, 总共 59 条记录 每页显示 5 条记录

1 2 3 4 5

fdcs.elab.cnic.cn:3000/dcmap

下载轨迹

名称: [输入框]

记录人: [输入框]

开始时间: 2015年8月11日 GMT+8 上午 12:15:02

结束时间: 2015年8月11日 GMT+8 上午 12:15:02

轨迹长度: 208114米

关键地点: [输入框]

文件大小: 18.63 MB



Next...

- ❖ 提供针对大数据，特别是跨学科融合大数据的统计分析工具，方便用户在平台上直接对采集的观测数据进行后续分析；
- ❖ 设计一种数据发布模式，为用户提供双向且自由的信息交流服务，确保在数据生产者拥有知识产权的前提下便于规范化的数据下载与引用；
- ❖ 研究富媒体数据的存取方法，将现代生态系统观测中，高频传感器与富媒体传感器自动获取的数据集成进平台当中；
- ❖ 数据可视化、可视化分析.....

谢谢

- ❖ 请联系我们:
- ❖ Email: chencan@cnic.cn
- ❖ Tel: 01058812709
- ❖ <http://fdcs.elab.cnic.cn:3000>