

数据共享平台的数据访问及其在 Android 终端的呈现^①

赵会敏, 雒江涛

(新一代信息网络与终端市协同创新中心, 重庆 400065)

(重庆邮电大学 电子信息与网络工程研究院, 重庆 400065)

摘 要: 实现数据共享, 可以使更多的人更充分地利用已有数据资源, 创造新的价值. CKAN(Comprehensive Knowledge Archive Network)是世界领先的开源数据平台, 对收集的数据实现管理和共享. 本文实现了一款基于 Android 开发的 CKAN APP, 以使用户可以随时随地查看 CKAN 平台上的共享数据. 本文主要是运用 Android 客户端的开发技术, 通过 CKAN 的 API 来访问 CKAN 平台的数据资源, 然后将访问到的数据资源转化为合适的格式存储起来, 最终将数据在 Android 终端呈现出来. 测试结果表明: 本文开发的 CKAN APP, 为用户随时随地查看 CKAN 平台的共享数据提供了一个可用的 APP 工具.

关键词: 数据共享平台 CKAN; Android 终端开发; CKAN APP; 数据资源的访问; 终端呈现

Data Access to Data Sharing Platform and Presentation in Android Terminal

ZHAO Hui-Min, LUO Jiang-Tao

(Collaborative Innovation Center for Information Communication Technology, Chongqing 400065, China)

(Electronic Information and Networking Research Institute, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Realizing data sharing can make more people take full advantage of the existing data resources to create new value. CKAN (Comprehensive Knowledge Archive Network) is the world's leading open source data platform. It can manage the existing data resources and make them sharing. So this paper implements a CKAN APP based on Android development in order that the users can view the sharing data resources of CKAN anytime and anywhere. This paper mainly uses the Android client development technology to realize the access into data resources of CKAN by CKAN API, then changes the format of the obtained data resources into an appropriate format and stores them, eventually, displays the data resources at the Android terminal. The test results show that CKAN APP developed in this paper provides an available APP tool for users to view the sharing data resources of CKAN anytime and anywhere.

Key words: data sharing platform CKAN; Android terminal development; CKAN APP; data resource access; terminal presentation

随着信息时代的不断发展, 不同部门、不同地区间的信息交流逐步增加, 计算机网络技术的发展为信息传输提供了保障. 当大量的空间数据出现在网络上, 由于不同的用户提供的数据来自不同的途径, 其数据内容、数据格式和数据资源千差万别, 因此给数据共享带来了很大困难. 数据共享就是让在不同地方使用不同计算机、不同软件的用户能够读取他人数据进行

各种操作运算和分析. 实现数据共享, 可以是更多的人更充分地使用已有数据资源, 减少资源收集、数据采集等重复劳动和相应费用^[1]. 大数据时代的到来, 开放数据、实现数据的共享使得更多参与者能对数据进行加工挖掘, 从而创造新的价值.

CKAN 是一个非常强有力的数据共享平台的数据管理系统. 是由英国一个非营利性组织 —— 开放知

① 基金项目:重庆市应用开发计划项目(cstc2013yykfA40006)

收稿时间:2015-07-21;收到修改稿时间:2015-09-08

识基金会(Open Knowledge Foundation)开发并用于网站 Datahub.io 的使用. 它通过提供发布数据、分享数据、查找数据和使用数据的功能, 让数据更加容易使用和处理. CKAN 目的在于为那些想要使他们收集到的数据开放并实现共享的数据发布者(例如国家政府、地区政府、公司、组织)提供服务. 数据只要被发布, 用户就可以利用其关键字或筛选标签来搜寻他们需要的数据, 并且可以用表格、曲线图和地图示意图(提供在地图上按地区搜索地理空间数据的功能)三种方式来预览数据^[2]. 现在, CKAN 已经为全世界超过 40 个数据中心所使用, 其中有门户网站、国家政府(如英国的 data.gov.uk、欧盟的 publicdata.eu), 他们用 CKAN 来发布官方、社区网站的数据^[3].

在我国互联网的发展过程中, PC 互联网已日趋饱和. 然而, 随着近几年手机、平板电脑等智能移动终端的出现和普及, 以及 3G、4G 的发展和 WIFI 的广泛铺设, 移动互联网呈现井喷式发展. APP 作为移动互联网的主要入口, 已被人们广泛运用^[4]. 由于移动智能终端的便携性, 用户可以在有网络的地方随时运用 APP 查看新的消息, 享受网络的服务. 移动 APP 给人们带来了极大的方便和乐趣^[5], 如果有一款实现数据共享的 APP, 那么需要数据的用户就可以随时随地搜寻数据、预览数据. 然而, 经过调研并未发现有这样功能的 APP^[6,7]. 基于此, 本文实现了一款基于 Android 的 CKAN APP 的开发.

1 CKAN APP的实现

本文基于 Android 的 CKAN APP 开发的关键技术思路是运用 Android 客户端的开发技术, 通过 CKAN 的 API 来访问 CKAN 平台的数据资源, 然后将访问到的数据资源转化为合适的格式存储起来, 最终将数据在 Android 终端呈现出来.

实现总体框图如图 1 所示.

1.1 搭建自己的 CKAN 平台

要运用 CKAN 实现数据共享, 需要先搭建自己的 CKAN 平台. CKAN 的安装方式有三种^[8]: 用操作系统包安装, 用源的方式安装以及用 Docker 镜像安装. 本文选择了用最简单、最快速的用操作系统包安装的方式^[9], 此种安装方式需要 Ubuntu 12.04, 64-bit 的操作系统. 本文搭建的 CKAN 网站作为后续 Android 开发时访问的服务器使用, 所以我们并不采用把 CKAN 单

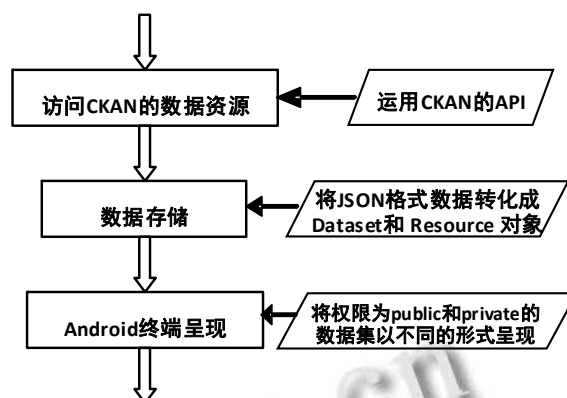


图 1 实现总体框图

独安装在一台电脑上只作为服务器, 而用另一台电脑做 Android 开发的方法. 我们用同一台电脑运行 CKAN 并且开发 Android. Android 的开发需要的操作系统是 Windows, 而运行 CKAN 需要 Linux 系统, 因此我们需要借助虚拟机来安装 Ubuntu 12.04, 64-bit 操作系统, 并在此基础上搭建 CKAN 平台.

1.2 访问 CKAN 数据集

网络连接的成功与否直接决定了此 APP 能否正常使用, 因此首先必须保持网络能够正常连接. 通过 API URL 来访问 CKAN 的平台的数据集, 大体可以分为两种方式, 一种是不标明版本号(如 http://demo.ckan.org/api/action/package_list), 另一种则是标明版本号(如 http://demo.ckan.org/api/3/action/package_list). 不标明版本号时, CKAN 便自动选择最新版本的 CKAN API, 标明版本号时, CKAN 便选择相应版本的 API. 因为 CKAN 会不断升级更行, 考虑到客户端使用的 API 版本与接受 API URL 请求的 CKAN 平台的 API 版本相同问题, 即版本的兼容性问题, 本文采用标明版本号的方式, 使用 CKAN API V3^[10].

为了获取 CKAN 网站的数据集及相应数据集里面的资源以及其它信息, 使用 API: `ckan.logic.action.get.current_package_list_with_resources`, 发送相应 HttpPost 请求, 发送前对该 HttpPost 加入一个含有 `SysAdmin-UserAPIKey` 的 Header, 只有获 `SysAdminUserAPIKey` 这个字段, 才有权限访问这个网站的数据. 发送请求后, 服务器经过判断之后做出响应, 之后返回该网站的数据集信息. 为了获取某个网站特定用户的相关信息和某个组织里面的人员以及其他相关信息, 分别使用 API: `ckan.logic.action.get.user_show` 和 `ckan.logic.action`.

get.organization_show, 利用以上相同的思路获取。

在每次准备与服务器进行通信之前, 我们都会通过函数 isNetAccessible() 来判定现在是否可以与服务器对话^[11]。需要它的原因有如下几个: 测试本机现在是否已打开网路连接(连接 WIFI 或使用移动数据网络上); 阻止过于频繁地向服务器发送请求, 从而减轻服务器的负担; 当 Android 客户端与服务器通信时, 不能向服务器发送请求, 为了防止再次向服务器发送请求, Android 客户端实现时设置一个 isRequestRunning 变量, 这个变量为 boolean 值, 用这个变量来判断 Android 是否在与服务器交换数据中。

1.3 数据存储

由于我们要将从 CKAN 网站获取的数据集及其数据资源在 Android 终端显示出来, 但 Android 或 Java 没有自带的相关的类来描述 CKAN 平台数据集及其资源的类, 所以为了存储从 CKAN 网站获得数据, 我们需要定义两个新的类: Dataset 和 Resource。

Dataset 的成员数据如下: private String title; private String id; private ArrayList<Resource> resources。

title 用于存储该数据集的名字, id 用于存储该数据集的特有 id, 用于区别两个名字不同的数据集; resources 则存储该数据集的相关数据资源, 注意到一个数据集内可以有零个或多个资源, 所以我们用 ArrayList<Resource> 的形式来存储。

Resource 的成员数据如下: private String name; private String url; private String id; private String datasetId。

name 用于表示该资源的名字, url 用于表述该资源在服务器上面的统一资源定位符, id 用于存储该资源的特有 id, 用于区别两个名字不同的资源, datasetId 用于表述该资源所属数据集 ID。

在收到服务器发送给 Android 客户端的数据后, 该数据是以 JSON 格式被传送的, 不便于 Android 端的使用。所以需要将收到的 JSON 数据中自己需要的数据提取出来, 并初始化成对应的 Dataset、Resource 对象, 并保存在 Android 端本地, 供随时使用。完成这个功能的是 JsonUtil 类, 专门用来处理收到的 JSON 数据。

1.4 数据集在 Android 终端的呈现

因为 CKAN 的数据集的权限有的是公开的(public), 有的是非公开的(private), 非公开的数据只有

是在一个组织里面的用户才有权限浏览。因此, 当数据集在 Android 终端呈现时, 我们采用不同的方式来处理这两种数据集。把所有的 public 数据集的信息放入 showDatasets 对象中, 由于显示数据集时正是将 showDatasets 对象里面的数据进行显示的, 所以在登入网站时, 就将所有的 public 数据集显示出来。而对于 private 数据集, 我们需要先用数据所在组织的用户名去搜索, 此时直接把实例化的 userDatasets 对象的所有数据集放入到 showDatasets 对象中, 然后呈现在屏幕上。采用通过 Android 的 WebView 来进行显示, 这种方式可以显示所有格式的资源, 并且可以将 csv 格式的资源内容以表格和曲线图的形式显现^[12]。

2 运行与测试

测试分为两部分, 一个是 CKAN 平台的测试, 一个是对 CKAN APP 的测试。

2.1 对 CKAN 平台的测试

通过向 CKAN 平台上传数据及预览数据, 来测试 CKAN 平台是否真正搭建成功, 然后将 CKAN 平台的网址发布, 以便所用用户都可以访问此平台, 经过网络设置我们搭建的 CKAN 平台可以用网址 http://113.251.164.176:7777/ 进行访问。至此此 CKAN 平台可以作为一个 Server, 供 Android 客户端访问使用。

2.2 对 CKAN APP 的测试

在 Android 上安装好 CKAN APP 后, 我们点击 CKAN 图标, 进入 APP, 访问的是我们自己搭建的 CKAN 平台, 因为 public 数据集和 private 数据集的呈现方式不同, 所以登入网站首先呈现的是所有的 public 数据集, 结果如图 2 所示。



图 2 public 数据集呈现图

当我们查询某个组织的 private 数据集时, 需要用

这个组织的用户名查询, 查询结果如图 3 所示.

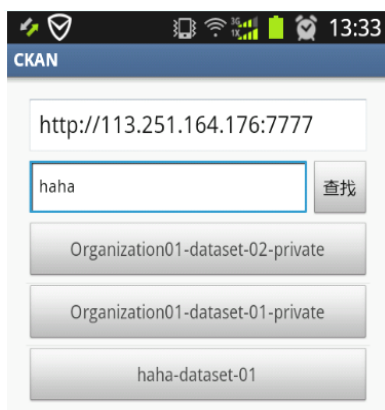


图 3 private 数据集的呈现图

此 APP 可以查询所有的数据格式, 在此展示一下对 CSV 格式文件的两种查看方式, 图 4 以表格(Grid)的方式查看、图 5 以曲线图(Graph)的方式查看.

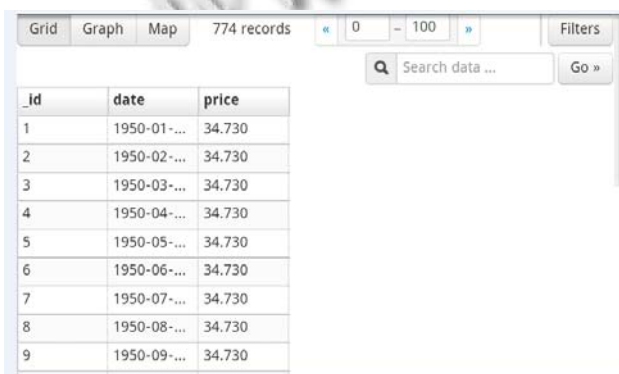


图 4 以 Grid 的方式查看 CSV 格式的文件

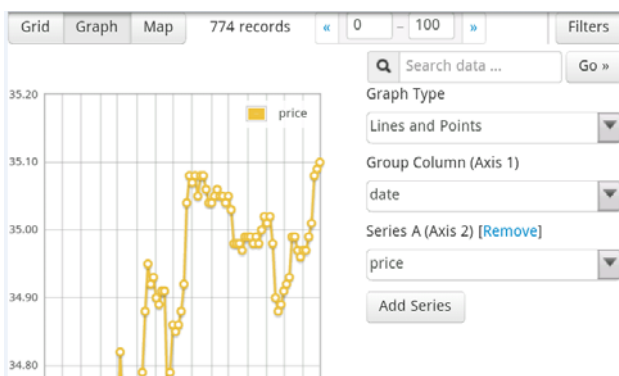


图 5 以 Graph 的方式查看 CSV 格式的文件

3 结语

测试表明, 本文完成了基于 Android 的 CKAN APP 的开发, 达到了我们最初的目标: 用户可以随时随地利用基于 Android 的 CKAN APP, 查看 CKAN 平台上的共享数据, 查询用户需要的数据, 充分利用共享数据. 此 APP 可以查寻和访问 CKAN 平台上的所有资源, 但还未能实现下载, 上传数据等功能, 这也是受到手机存储能力限制的结果. 此外, 没有对界面进行专业的美化. 这将会是我们以后努力的方向, 争取开发出一款功能更齐全, 界面更美观的 CKAN APP.

参考文献

- 1 张英俊, 谢斌红, 郭勇义. 元数据技术在科学数据共享平台中的应用. 太原理工大学学报, 2009, 40(4): 341-344.
- 2 <http://ckan.org/>. [2015-04-09].
- 3 <http://ckan.org/instances/>. [2015-04-10].
- 4 肖志辉. 移动互联网研究综述. 电信科学, 2009, 25(10): 30-36.
- 5 Shneiderman B, Plaisant C, 李健利, 张国印. 用户界面设计: 有效的人机交互策略. 第 5 版. 北京: 电子工业出版社, 2009.
- 6 <http://shouji.baidu.com/>. [2015-04-25].
- 7 <http://app.mi.com/>. [2015-04-25].
- 8 <http://docs.ckan.org/en/latest/maintaining/installing/index.html/>. [2015-04-29].
- 9 <http://docs.ckan.org/en/latest/maintaining/installing/install-from-package.html>. [2015-04-29].
- 10 <http://docs.ckan.org/en/latest/api/index.html?highlight=api/>. [2015-05-04].
- 11 Eckel B. Thinking in java 4. 4th ed., Ametica: Prentice Hall PTR, 2007: 193-273.
- 12 <http://android.toolib.net/reference/android/content/pm/ActivityInfo.html/>. [2015-06-10].