



Microsoft
Open Technologies

Open Data Platform for China Government

– Open Source, Open Standards and Open Data

Emily Chen

emilychen522@gmail.com

emchen@Microsoft.com

I am

GNOME.Asia Founder

GNOME Foundation Board 2010 – 2011

Beijing GNOME Users Group Founder

Founder Member of Kaiyuanshe

Senior Software Engineer at SUN/Oracle 2006 -2014

Program Manager of Microsoft Open Technologies



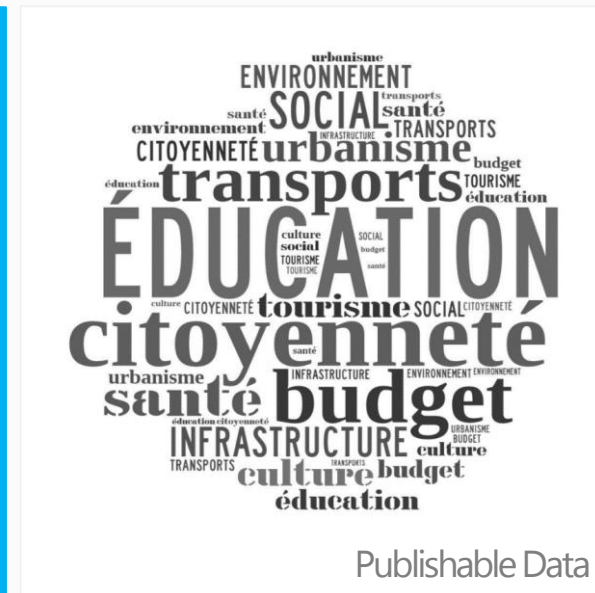
Agenda

- Background of Government Open Data
- Open Data Portal (CKAN) + Open Standard (OData)
- Open Data Platform Architectures and Functions
- Case Study
- References



Open Data Defined

Making public data broadly accessible and usable by humans and machines, free of any technological, legal or usability barriers



A movement that benefits all actors in the modern society and economy

Open Data Around the World

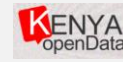
“Release of public sector information creates opportunities for innovative use and reuse of data, and allows the commercial, research, and community sectors to add value.”

Australian Government [2013]



“It enables data-driven decision making: parliamentarians, policy makers, civil society organizations and individuals can see progress and make accurate, informed decisions on issues that affect people's lives.”

Kenya Open Data Initiative [2013]



“Openness will strengthen our democracy and promote efficiency and effectiveness in Government.”

Barack Obama, President of the United States [January 2009]



“The next generation of innovation in... academic research or public services will be possible because we find new ways to harness and learn from data.”

World Economic Forum [December 2013]



“It makes the government more accountable to citizens and strengthens our democracy. It brings us better public services and it feeds economic and social growth.”

Data.gov.uk [2013]



“By opening up more data, and through innovative use of technology, we can crowd-source ideas and co-create applications with the wider community.”

Tharman Shanmugaratnam, Singapore Deputy Prime Minister [April 2013]



“Not only are we making more effective decisions through improved use of data, but also involving communities, making them more informed and empowering them to deal with future challenges.”

World Economic Forum [January 2013]



Open Source Data Portal Platform - CKAN

- OKFN CKAN

OKFN is a non-for-profit UK organization and develop the CKAN OSS product, a leading Open Data Catalogue and Indexing content management system running on Linux Ubuntu

- CKAN Open Data site examples:

- data.gov.uk UK government web site CKAN integrated with Drupal and Azure Datamarket hosting UK Weather opendata dataset
- USA data.gov with Socrata, Drupal

The image shows a screenshot of the CKAN website. At the top, the CKAN logo is displayed with the tagline 'The open source data portal software'. Below the logo, there's a navigation bar with links like 'Case studies', 'Feature Tour', 'Developers', 'FAQ', and 'Get In Touch'. The main content area features a large banner for 'UK: Adur District Council Spending' with a title 'UK: Adur Data' and a subtitle 'From Spikes Cavell, Spotlight'. To the right of the banner, there's a section titled 'CKAN, the world's leading open-source data portal platform' with a description of CKAN's purpose and three buttons: 'Take Tour', 'Get CKAN', and 'Live Demo'. Below the banner, there's a 'Feature Overview' section with icons and text describing various features like 'Complete catalog system', 'Strong integration with third-party CMS's', 'Data visualization and analytics', 'Workflow support', 'Fine-grained access control', 'Integrated data storage and full data API', and 'Federated structure'. On the far right, there's a 'Support and Hosted Solutions' section with text about CKAN being open source and available for free, and a link to 'Read More'.

ckan The open source data portal software

UK: Adur District Council Spending

demo.ckan.org/dataset/adur_district_spending

UK: Adur Data

From Spikes Cavell, Spotlight

For Adur, records from April

Data and Resources

Adur District Council

Adur District Council Ap

CKAN, the world's leading open-source data portal platform

CKAN is a powerful data management system that makes data **accessible** – by providing tools to streamline **publishing**, sharing, **finding** and **using** data. CKAN is aimed at data publishers (national and regional governments, companies and organizations) wanting to make their data open and available.

Take Tour Get CKAN Live Demo

Feature Overview

- Complete catalog system with easy to use web interface and a powerful API
- Strong integration with third-party CMS's like Drupal and WordPress
- Data visualization and analytics
- Workflow support lets departments or groups manage their own data
- Fine-grained access control
- Integrated data storage and full data API
- Federated structure: easily set up new instances with common search

Read More

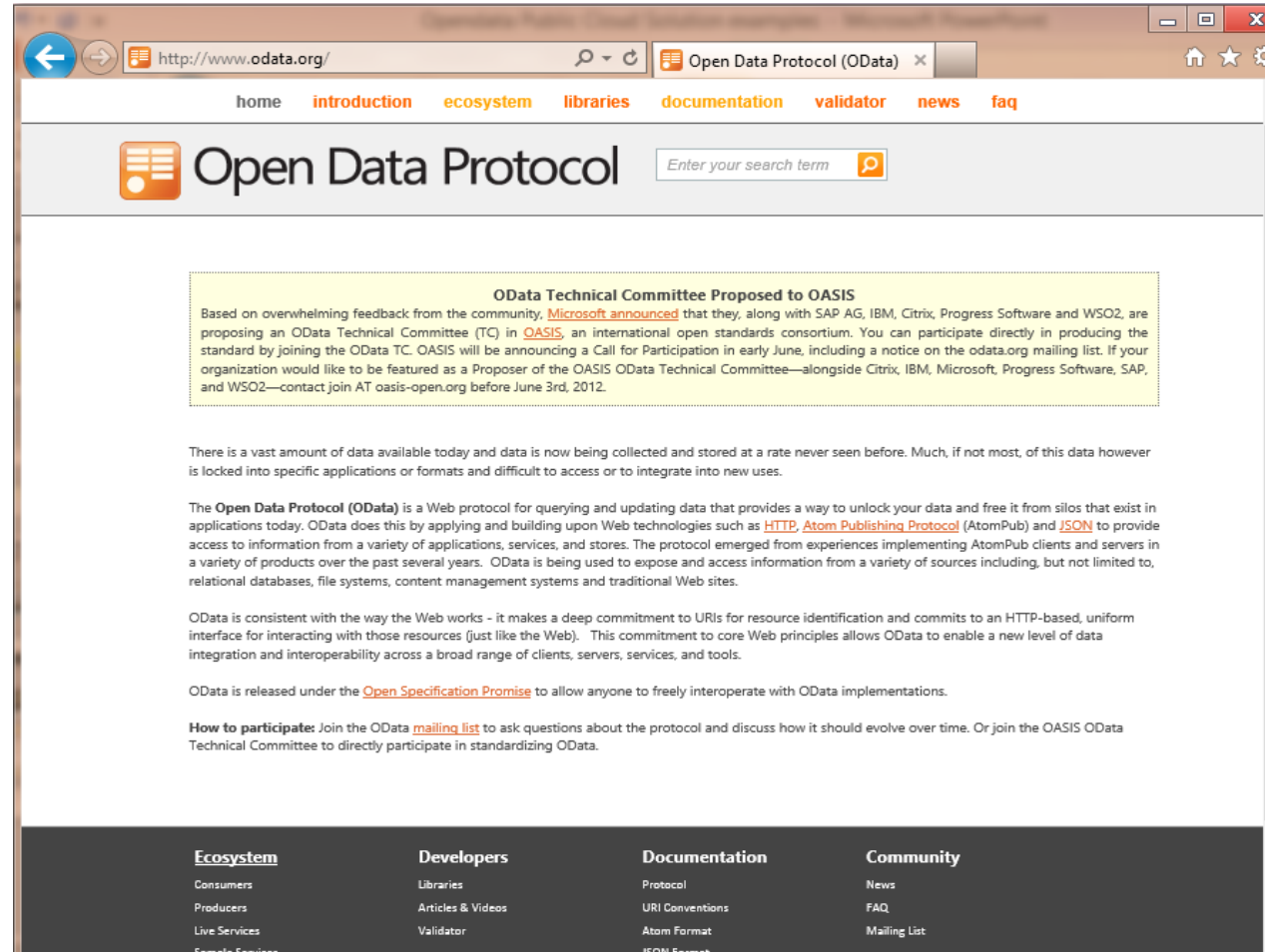
Support and Hosted Solutions

CKAN is open source and can be downloaded and used for free. Users can also get hosting and support from a range of suppliers.

A full-time professional development team at the Open Knowledge Foundation maintains CKAN and can provide full

Open Data – OData moving to Standards

- Opening data silos through OData protocol – multi-platform and multi-vendor
- Open Data protocol, built on web standards open spec moving to standardization
- OASIS Approves Odata 4.0 Submitted by Microsoft, IBM, SAP, Citrix, Progress Software and WSO2
- OData ecosystem and supported products/services



CKAN + OData = Open Data Platform

- **MS Open Tech's Contribute to OSS project**

- OData plugin for CKAN
- OData feeds generator for JDBC
- OData Producer for SugarCRM
- CKAN Images on VM Depot

- **Business project**

- Open Data Project for Chinese Government
- Powered on Azure Public Cloud

- **Ecosystem**

- Juhe Dataset's API support OData
- Odata.org community
- National Open Data Contest

The screenshot displays the VM Depot website, which is a platform for finding, deploying, and sharing Open Source Software (OSS) images. The header includes the VM Depot logo and navigation links for '订阅' (Subscribe), '登录' (Login), and '加入社区' (Join Community). The main content area features a search bar with 'CKAN' entered, and a list of virtual machine images. The first image is for CKAN, published by '刘铮' (Liu Zheng) on the 'ubuntu' platform. It includes a description in Chinese, a list of software packages (apache, ckan, nginx, postgresql, pylons, solr), and a tag 'odata'. The bottom section of the image shows the OData.org website with the title 'Open Data Solution Using CKAN and OData' and a brief description of the solution.

VM Depot PREVIEW
by Microsoft Open Technologies

订阅 登录 并加入社区 中文(简体)

返回 创建虚拟机 部署脚本 发布 查看 共享 帮助

浏览镜像 我的账户

虚拟机镜像

VM Depot - 为 Windows Azure 查找、部署并共享 OSS 镜像

VM Depot 是可以轻松部署在 Windows Azure 上的预配置操作系统、应用程序和开发栈的社区主导型目录。查找您喜爱的软件并快速部署，或加入社区、建立虚拟机镜像然后与他人分享。了解更多。

CKAN

特性 添加日期 名字 平台 评级

CKAN

本镜像一键部署整套CKAN环境到微软公有云 Azure，并预装了常用插件：datastore, datapusher, odata, hierarchy, odctheme, resource_proxy. 【注意：1. 虚拟机创建完毕后，需修

发布者：刘铮
平台：ubuntu

软件包：apache ckan nginx postgresql pylons solr
标签：ckan odata

创建虚拟机 部署脚本

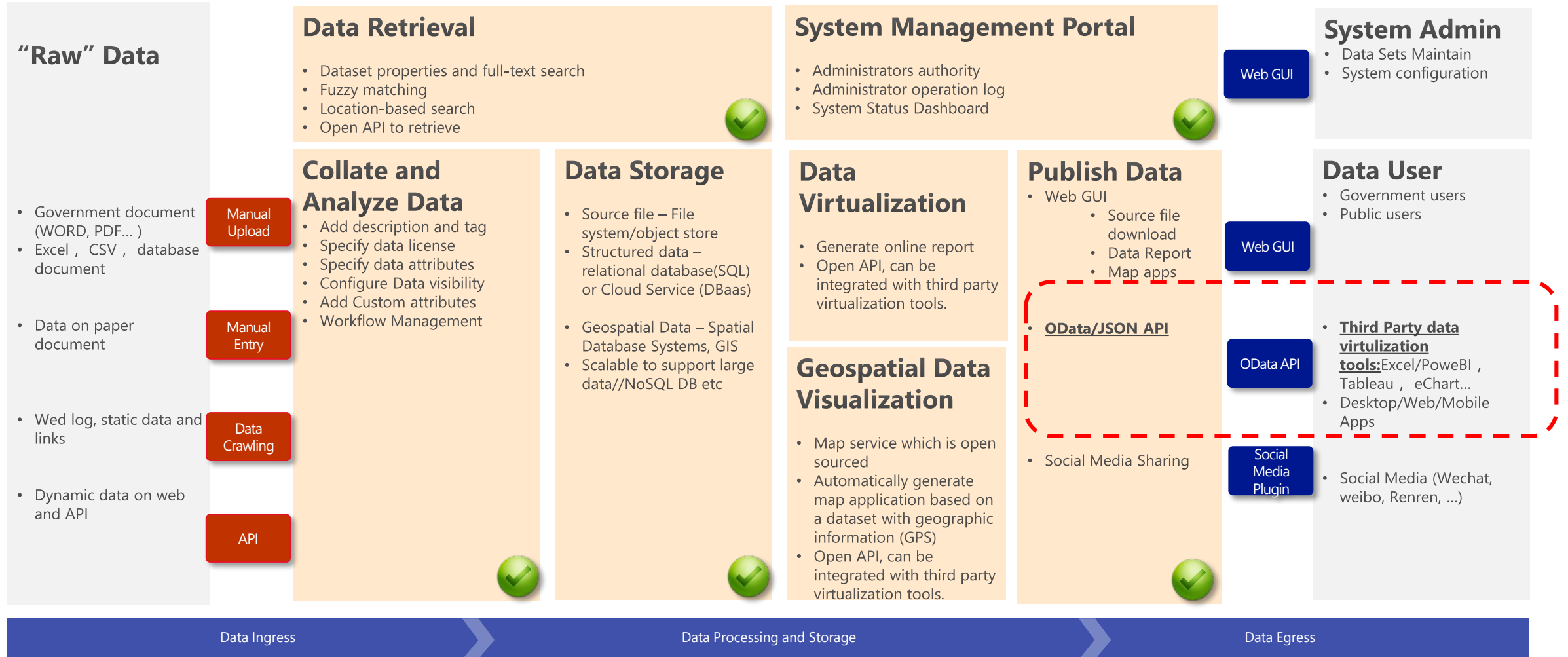
www.odata.org

OData Home Blog Getting Started Documentation Libraries

Open Data Solution Using CKAN and OData

Today, MS Open Tech China released a new, open data solution for Microsoft Azure built around CKAN, an open source data portal, and an optimized OData plugin developed to specific requirements for the Chinese market.

Open Data Platform Architecture



Support
Open API

Data Import – Process Government Data, Document and links

- Three steps: Create Data Sets -> Import Data -> Configure Data

This screenshot shows the first step of the data import process: '1 创建数据集' (Create Dataset). The interface includes a title field with the text 'xxx市A级餐馆地址和联系方式', a URL field with 'xixiankan.chinacloudapp.cn/dataset/xxx-a', and a description field with 'xxx市A级餐馆地址和联系方式, 包含GPS经纬度数据'. There are also fields for tags (餐饮, 地理信息, 食品卫生, 联系方式), a license dropdown (其他(开放)), an institution dropdown (1xxxq), and a visibility dropdown (公开). A '下一步: 添加数据' (Next Step: Add Data) button is at the bottom right.

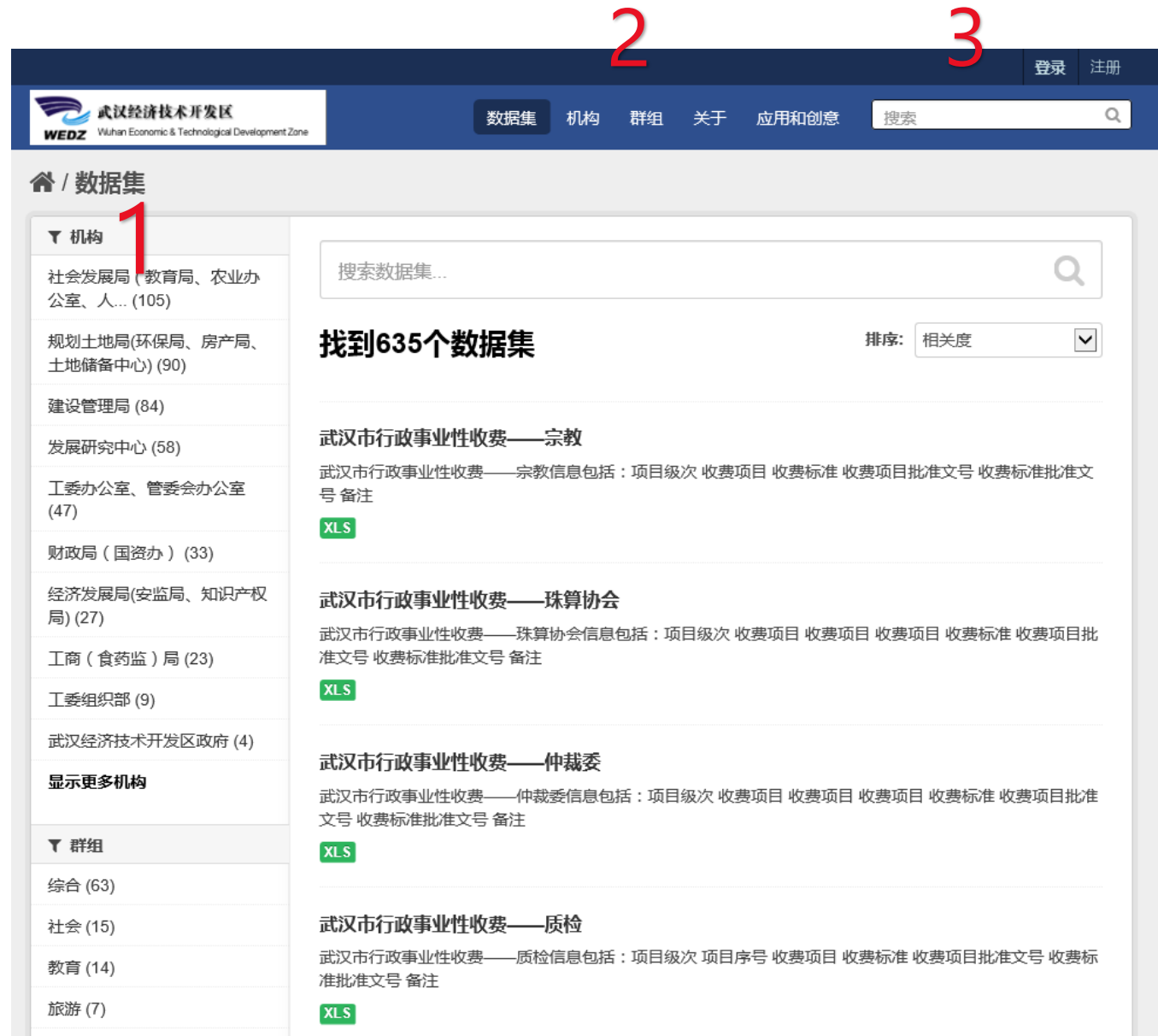
This screenshot shows the second step of the data import process: '2 添加数据' (Add Data). It features a '文件:' (File) section with '上传' (Upload) and '链接' (Link) buttons. The '名称:' (Name) field contains 'XX市A级餐馆信息', and the '描述:' (Description) field contains 'XX市A级餐馆信息, 包含联系方式和GPS经纬度坐标'. A '格式:' (Format) dropdown is set to 'CSV'. Navigation buttons '上一步' (Previous Step) and '保存并添加另一个' (Save and Add Another) are at the bottom.

This screenshot shows the third step of the data import process: '3 额外信息' (Additional Information). It contains fields for '源:' (Source) with 'XX市A级餐馆信息', '版本:' (Version) with '2015年3月, 第一版', '作者:' (Author) with 'XXXX', '作者邮箱:' (Author Email) with 'xxxx@xx-city.gov.cn', '维护者:' (Maintainer) with 'YYYY', and '维护者邮箱:' (Maintainer Email) with 'yyyy@xx-city.gov.cn'. There are also three '自定义字段:' (Custom Fields) sections, each with 'Key' and 'Value' input fields. The first custom field has '自定义key1' and '2015'. Navigation buttons '上一步' (Previous Step) and '完成' (Complete) are at the bottom right.

- Support multiple document format– WORD, EXCEL, CSV, JPG, HTML, URL...

Data Collate, Category and Search

1. Data Sets Query in different Category –Organization , Groups, Data types, attributions...
2. Three view points :Data Sets, Organization and Groups
3. Full text search and fuzzy matching



The screenshot displays the WEDZ Data Set Query interface. The top header (labeled '2') includes the WEDZ logo, navigation links (数据集, 机构, 群组, 关于, 应用和创意), a search bar, and login/register buttons. The left sidebar (labeled '1') shows a tree view of data sets categorized by Organization (机构) and Groups (群组). The main content area (labeled '3') shows search results for '武汉市行政事业性收费' (Wuhan Administrative and Public Service Fees), displaying a list of fees with columns for item level, fee item, fee standard, fee item approval number, fee standard approval number, and remarks. Each entry has an 'XLS' download button.

机构

- 社会发展局 (教育局、农业办公室、人... (105)
- 规划土地局(环保局、房产局、土地储备中心) (90)
- 建设管理局 (84)
- 发展研究中心 (58)
- 工委办公室、管委会办公室 (47)
- 财政局 (国资办) (33)
- 经济发展局(安监局、知识产权局) (27)
- 工商 (食药监) 局 (23)
- 工委组织部 (9)
- 武汉经济技术开发区政府 (4)
- 显示更多机构

群组

- 综合 (63)
- 社会 (15)
- 教育 (14)
- 旅游 (7)

搜索数据集...

找到635个数据集

排序: 相关性

武汉市行政事业性收费——宗教

武汉市行政事业性收费——宗教信息包括：项目级次 收费项目 收费标准 收费项目批准文号 收费标准批准文号 备注

XLS

武汉市行政事业性收费——珠算协会

武汉市行政事业性收费——珠算协会信息包括：项目级次 收费项目 收费项目 收费项目 收费标准 收费项目批准文号 收费标准批准文号 备注

XLS

武汉市行政事业性收费——仲裁委

武汉市行政事业性收费——仲裁委信息包括：项目级次 收费项目 收费项目 收费项目 收费标准 收费项目批准文号 收费标准批准文号 备注

XLS

武汉市行政事业性收费——质检

武汉市行政事业性收费——质检信息包括：项目级次 项目序号 收费项目 收费标准 收费项目批准文号 收费标准批准文号 备注

XLS

Data Virtualization – GIS as example

- When Data entry contains geographic information(GIS), like latitude and longitude, open data platform will automatically generate a map application
- Open Data Platform support basic Charting tools

武汉市A级餐饮信息

网址 <http://xixiankan.chinacloudapp.cn/dataset/2cad2a95-5cbf-4b1f-973f-b2dbcaa91ffa/resource/9fbb6681-77ce-44c8-a1d1-4a1202760b3e/...>

武汉市A级餐饮信息

Grid Graph Map 15 records 0 - 100 Search data ... Go » Filters

	Name	Address	Street	Tel	Status	Type	Lan	Long
	武汉天外天莱花园度...	武汉经济技术开发区军山街...	军山街		正常	大型餐馆	114.11604...	30.39236...
	武汉迈森豪餐饮管理有...	武汉经济技术开发区龙王工...	沌口街	84286437.0	正常	企业事业机...	114.12998...	30.46472...
	武汉经济技术开发区江...	武汉经济技术开发区江汉大学	沌口街	84225285.0	正常	高等院校...	114.16081...	30.50999...
	上海怡乐食品科技服...	武汉经济技术开发区车城南...	沌口街	64263615	正常	企业事业机...	114.14831...	30.48070...
	武汉观湖酒号酒店管理...	武汉经济技术开发区观湖路9号	沌口街		正常	大型餐馆	114.15051...	30.49777...
	武汉湖锦酒楼管理有限...	武汉经济技术开发区博学路2...	沌口街	85667777	正常	特大型餐馆	114.16076...	30.50261...
	武汉一撮毛美食有限责...	武汉经济技术开发区沌口街...	沌口街		正常	大型餐馆	114.10829...	30.45991...
	武汉惠远经济发展有限...	武汉经济技术开发区26M地块	沌口街		正常	特大型餐馆	114.15693...	30.49808...
	湖北明星酒店有限责任...	武汉经济技术开发区东风大...	沌口街	84478502	正常	大型餐馆	114.17253...	30.50611...
	武汉喜瑞得投资管理有...	武汉经济技术开发区3C2地...	经济技术...	027-8425...	正常	大型餐馆	114.15943...	30.49503...
	武汉实验外国语学校初...	武汉经济技术开发区龙阳大...	沌口街		正常	中小学校...	114.18034...	30.51686...
	武汉经济技术开发区赛...	武汉经济技术开发区嘎纳印...	沌口街		正常	托幼机构...	114.18196...	30.52002...
	武汉沌口长江大酒店有...	武汉市经济技术开发区创业...	沌口街		正常	大型餐馆	114.15704...	30.48706...
	武汉华鼎后勤配套服务...	武汉经济技术开发区神龙大...	沌口街	15337148...	正常	企业事业机...	114.17308...	30.49095...
	湖北明星酒店有限责任...	武汉经济技术开发区湘隆时...	沌口街		正常	大型餐馆	114.16775...	30.49901...



Case Study in China

Government Open Data Platform in China



武汉经济技术开发区
WEDZ
Wuhan Economic & Technological Development Zone

数据集 机构 群组 关于 应用和创意 搜索

搜索你的数据

比如:黄金价格

热门标签 2014 采购招标 收费

开放数据作品展示

特色应用程序

武汉开发区餐饮分布图
展现武汉开发区内A、B、C级餐厅分布情况
开发者: 微软开放技术
License: CC

武汉市2014年5月空气质量变化图
展现2014年5月份武汉市空气质量的变化
开发者: 微软开放技术
License: CC

武汉经济技术开发区2014年经济统计数据
可视化展示经济统计数据
开发者: 微软开放技术
License: MIT

武汉市日空气质量监测
实时显示武汉市日空气质量情况
开发者: BestApp
License: MIT

武汉市政务公开发展历程
可视化展现武汉市政务公开10年来的历程
开发者: 徐宝宏
License: MIT

武汉市二氧化碳排放情况
我们的排放地图http://www.whedz.cn...
开发者: 李健伟
License: MIT

References

1. Odata.org
2. CKAN.org
3. OData Plugin for CKAN in GitHub - <https://github.com/msopentechcn>
4. CKAN Images on VM Depot China - <https://vmdepot.msopentech.cn>

Our Team in MS Open Tech China:

Standards: Ling Chen <lingche@microsoft.com>

Engineering: Ming Liu <mingli@microsoft.com>

Open Source: Emily Chen <emchen@microsoft.com>