

# 美国 openFDA 数据公开对我国食品药品监管数据管理的启示

宗欣, 王迎利 (国家食品药品监督管理总局信息中心, 北京 100053)

**摘要** 目的: 分析美国食品药品监督管理局 (FDA) 公共数据开放项目 (openFDA) 的启动背景、运行模式和管理经验, 为我国食品药品监管数据管理工作提供参考。方法: 采用文献研究的方法, 对 FDA 政府网站有关专栏和 openFDA 项目网站以及在中国知网等数据库中检索到的文献资料进行分析研究。结果与结论: openFDA 项目自开展以来积累了大量数据, 采取提供应用程序接口、提供数据资料直接下载, 以及建立开发者社区为其提供交流和分享的平台多种方式进行公共数据开放, 并已取得了一定成效, 值得思考借鉴。

**关键词:** 大数据; 食品药品监管; 数据管理; 数据公开

中图分类号: R95 文献标识码: A 文章编号: 1002-7777(2017)09-0976-04

doi:10.16153/j.1002-7777.2017.09.002

## On OpenFDA and Enlightenments for Data Management of Food and Drug Administration in China

Zong Xin, Wang Yingli (Information Center of China Food and Drug Administration, Beijing 100053, China)

**Abstract Objective:** To analyze the background, mode and management experience of openFDA, in order to provide references for the data management of food and drug administration in China. **Methods:** Relevant columns on FDA websites, openFDA project website, literatures found in the China National Knowledge Internet (CNKI) and other databases were analyzed using the method of literature analysis. **Results and Conclusion:** The openFDA project has accumulated a lot of data since its implementation and has provided several to improve the data opening of openFDA, including providing application interface, direct data download as well as establishing, a developer community to provide communication and sharing platform, and has been proved effective in practice, which is enlightenments for us.

**Keywords:** big data; food and drug administration; data management; data opening

### 1 openFDA公开数据简介

#### 1.1 背景

2012年3月29日, 美国奥巴马政府提出了“大数据研发计划”(Big Data Research and Development Initiative)<sup>[1]</sup>; 2014年6月2日, 美国食品药品监督管理局公共数据开放项目(openFDA)正式

启动上线, 这开启了美国政府公众健康数据库向广大网络开发人员、研究者和民众开放的大门, 以用户友好的方式向民众提供了海量的数据资源, 使开发人员、研究人员和公众能够更加便捷地访问和挖掘利用FDA所积聚起的、大量的医疗健康数据<sup>[2-3]</sup>。

在中国,自2014年3月“大数据”首次写入《政府工作报告》以来,李克强总理多次提及并反复强调要开发应用好大数据这一基础性战略资源。可见,大数据已经上升为我国国家层面的战略,具体到医药领域,在中国大数据快速发展的背景下,参考openFDA的运行模式和经验,对提升中国医药

数据管理工作具有一定的借鉴意义。

## 1.2 openFDA数据平台模块介绍

openFDA数据开放平台主要分为openFDA简介、API学习、API端点、社区交流四个模块,见图1<sup>[4]</sup>。

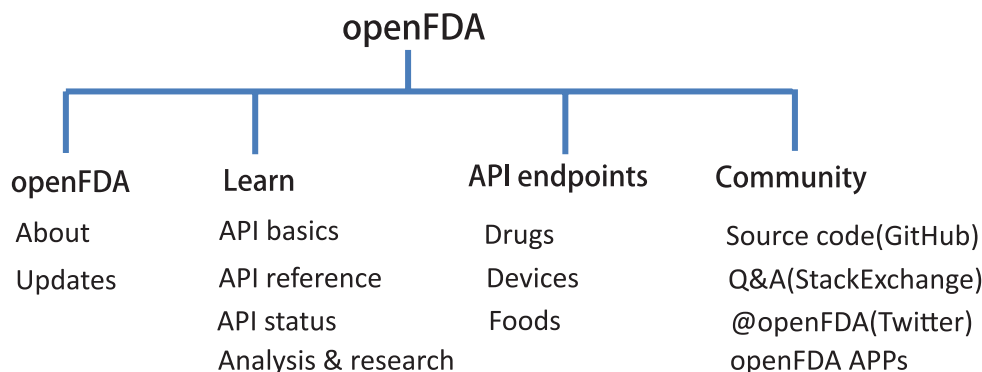


图1 openFDA数据平台模块

## 2 openFDA数据开放方式

openFDA允许技术专家、移动应用开发者、网络开发者、数据可视化专家以及研究人员更快地搜索、查阅和调用FDA数据库中的公共数据,并向民众即时推送大量来自FDA数据库的健康信息<sup>[5]</sup>。它提供了一个“基于搜索”的应用程序编程接口(API),允许开发人员能够像搜索引擎一样对数据库内的文本进行检索<sup>[6]</sup>。

openFDA还将开放的源代码和文档共享在GitHub(一个面向开源软件及私有软件项目的托管平台)上,鼓励行业开发人员基于这些开放的源代码和文档开发自己的应用,同时也使开发者有更大的灵活性来自行决定搜索哪种类型的数据,以及如何向终端用户呈现这些数据<sup>[7]</sup>。开发者可以向终端用户提供任何他们想要的数据类型,并在公用平台上设计发布多种多样的计算机应用程序(APP),促使在同一平台上出现各种各样的应用<sup>[6]</sup>。

openFDA提供了开放的API、原始数据下载、技术文档和应用实例,甚至还为重要的公共数据集建立了开发者社区(open community)<sup>[8]</sup>。在openFDA数据下载部分,既可以提供总量数据打包下载,也可以提供分年度数据下载。在开发者社

区,开发者可分享经验、APP及设计实例,分享途径包括GitHub、StackExchange(问答网站)和Twitter(社交网站)<sup>[3]</sup>。

## 3 openFDA平台数据内容模块

### 3.1 openFDA平台数据内容模块简介

在openFDA平台中,目前公开的数据内容主要包括医疗器械、药品和食品三大领域<sup>[9]</sup>。其中,应广大公众的需求和呼声,在openFDA平台运行伊始,将2004–2013年的380万次药品不良事件作为公开的第一个数据集进行了率先开放,截至2016年9月底,该部分药品不良事件数据已经激增增至682万次<sup>[9]</sup>。此外,通过社区互动,openFDA还在根据用户的需求和反映不断调整或增加新的数据种类。

截至目前,在医疗器械领域中,已经包括了医疗器械不良事件、召回、器械分类、510(k)、召回执行报告、上市前审批、唯一识别码、注册和备案8类相关数据;在药品领域中,包括药品不良事件、药品标签、药品召回执行报告3类数据;在食品领域中,包括食品召回执行报告数据。利用这些数据,移动开发人员就可以为消费者开发搜索类智能手机APP。利用这款APP,消费者可以查询是否有其他人在服用了某种药物之后也经历了相同的

药物不良事件等情况。

### 3.2 openFDA已公开数据汇总

根据openFDA网站显示,截至2016年9月底,通过openFDA平台公开的数据大部分已更新至同年9月底,即openFDA的数据更新情况良好。经笔者统计,截至2016年9月底已公开数据类别和总量见表1<sup>[9]</sup>。

表1 openFDA已公开数据类型和数据总量

| 领域   | 分类      | 数据总量 / 条 |
|------|---------|----------|
| 医疗器械 | 不良事件    | 5625939  |
|      | 召回      | 84779    |
|      | 分类      | 6128     |
|      | 510 (k) | 144579   |
|      | 召回执行报告  | 11387    |
|      | 上市前审批   | 32833    |
|      | 唯一识别码   | 670387   |
|      | 注册和备案   | 225082   |
| 药品   | 不良事件    | 6828534  |
|      | 标签      | 96775    |
|      | 召回执行报告  | 6372     |
| 食品   | 召回执行报告  | 36454    |

注:根据风险等级不同,FDA将医疗器械分为三类;部分医疗器械在美国销售前需进行“产品上市登记”(PMN:Premarket Notification)认证;该登记认证提交的文件需满足联邦食品、药品与化妆品法案(FD&C Act)第510章节,故通常将“产品上市登记”认证称为510(k)认证<sup>[10]</sup>。

## 4 openFDA数据公开的作用、影响及启示

### 4.1 发挥政府部门数据资源优势 有效刺激创新

政府相关部门在获得数据方面具有一定的优势,政府对数据进行公开可以大大节约社会成本,拉动相关产业的创新和发展,尤其是数据分析行业,将会随着数据公开得到迅猛的发展。当前,我国食品药品监管数据公开仅限于数据查询,类似美国API查询等数据、源代码下载等服务还未向行业开发者和普通民众开放。借鉴美国openFDA的经验,建议可以考虑挑选某一类数据源代码和数据本

身进行试点公开,待运行成熟后可进一步拓展公开数据种类的范围,强化落实数据开放政策。

### 4.2 联合社会第三方力量 大大提高公开数据的应用性

openFDA计划的启动和推广实际上是FDA内部和外部开发者共同深入研究的结果<sup>[5]</sup>。美国FDA拥有各种各样的数据类型,从基因到监管,再到临床研究。在内部进行紧张筹划的同时,他们还邀请了硅谷Iodine公司共同进行研发,在采用基于搜索的API之前,开发人员和研究人员几乎不可能轻松访问到这些数据。例如:之前,如果想查询对乙酰氨基酚(acetaminophen),需要下载所有文档,将其整合后,还要删除重复内容,该过程可能需要几年时间。虽然之前数据也是开放的,但访问起来很困难<sup>[8]</sup>。采用基于搜索的API技术之后,获得数据的过程变得更加容易、更加透明了。借鉴openFDA经验,建议我国有关政府部门加强与一流的数据开发公司和技术开发公司合作,开发适合于中国的检索技术和方案,提供高效的数据公开服务。

### 4.3 建立社区式互动平台 深化数据公开和交流

openFDA不仅提供了开放的应用程序接口,支持对数据进行下载,还建立了开发者社区,在这一官方建立的社区内,开发者可以进行交流,并分享数据开发经验。这对促进数据公开工作不断完善、为开发者提供一个沟通和交流的平台大有裨益,有助于促进用户界面更加友好化,进一步减少重复性开发,也有助于及时有效解决数据公开和应用过程中发现的问题,塑造监管部门和数据管理部门的正面形象。因此,建议在我国食品药品监管数据公开工作中能够兼顾此点,为数据开发人员、技术人员、食品药品研究人员和公众提供一个技术交流、技术共享的平台和空间,有效促进数据公开、数据挖掘等各项工作健康良性发展,促进提高监管效率和药品安全性水平。

### 4.4 以数据公开引导行业直面存在的问题 推进学术研究

openFDA公开的数据种类包括药物不良事件、510(k)等12类数据,在一定程度上是公开了行业可能存在的问题和过失,然而解决问题的前提首先是要直面问题,openFDA已经勇敢地迈出了公开数据这一步,有助于促进行业更好地面对问题、解决问题,推动进一步的研究与实践工作。目前,已有

我国研究人员对openFDA数据库中的药品不良反应数据进行分析<sup>[11]</sup>。当前我国处于改革的关键时期,面临一些亟待解决的问题,建议进一步加快食品药行业信息化建设步伐,整合现有信息化体系,共享融通数据资源,为食品药品监管数据公开打好软硬件基础,以数据公开推动提升监管效能,促进行业发展。

#### 4.5 通过对数据和信息处理后再公开 最大限度保护数据隐私和安全

由于openFDA只是提供处理后的可公开数据,不包含患者个人身份数据或其他敏感信息,既将有价值的信息进行了公开,同时还最大限度地保护患者的个人隐私和信息安全。结合到国内,开展食品药品监管数据公开工作同样需要对患者隐私和监管对象敏感信息等予以充分考虑,对相关数据进行脱敏后再予公开。

#### 4.6 通过数据公开 普及药品知识

在openFDA的实践当中,通过对食品药品监管数据进行处理和公开,特别是经过了众多行业开发者的二次开发之后,普通消费者在服用某种药物产生某种不良反应后,可以通过二次开发的APP等查询到是否有其他人在服用相同药物之后也产生了相同或类似的不良反应症状。使人们能够更加方便准确地了解自身的健康状况,了解更多器械、药物、食品方面的常识,对缓解民众医疗和用药焦虑、提高民众卫生意识、改善医患关系等方面都大有裨益。当前,我国同样存在患者信息不对称的现状,通过食品药品监管数据公开项目,这种现状有望得到一定程度的改善。

### 5 结论

openFDA公共数据开放项目的特色:一是积累了大量数据,这为各类决策分析等数据利用奠定了深厚的基础;二是提供多种方式支持相关方使用数据,包括提供应用程序接口供行业开发者进行二次开发、提供数据资料直接下载,以及建立开发者社区提供交流和分享平台。FDA公共数据开放项目在发挥政府部门数据资源优势和刺激创新方面起到了

一定的积极作用,而为了顺利推进该项目的实施并取得良好效果,FDA采取了提供多种公共数据获取方式、联合社会第三方力量促进公开数据应用、建立社区式互动平台促进数据公开和用户交流等措施并已见成效,值得思考借鉴。

#### 参考文献:

- [1] Whitehouse. Obama Administration Unveils "Big Data" Initiative: Announces \$200 Million in New R&D Investments [EB/OL]. [2016-10-03]. <https://www.whitehouse.gov/sites/default/files/microsites/ostp/bigdatapressreleasefinal2.pdf>. 2012.
- [2] FDA. Open Government: FDA Data Sets [EB/OL]. [2016-10-03]. <https://www.fda.gov/aboutfda/transparency/opengovernment/default.htm>.
- [3] FDA. Open FDA: Innovative Initiative Opens Door to Wealth of FDA's Publicly Available Data [EB/OL]. [2016-10-03]. <https://open.fda.gov/update/openfda-innovative-initiative-opens-door-to-wealth-of-fda-publicly-available-data/>.
- [4] FDA. Getting Started with Open FDA [EB/OL]. [2016-10-03]. <https://open.fda.gov/api/reference/>.
- [5] 包旭. 美国OpenFDA公众健康项目介绍[J]. 中国执业药师, 2015, 12(10): 18-22.
- [6] FDA启动openFDA项目向公众开放健康数据库[EB/OL]. [2016-10-07]. <http://yao.dxy.cn/article/77205>.
- [7] FDA. openFDA [EB/OL]. [2016-10-07]. <https://open.fda.gov/>.
- [8] 郭凡礼. 开放大数据带动美国医疗革命[N]. 中国经营报, 2014-06-16(12).
- [9] FDA. OpenFDA/download [EB/OL]. [2016-10-12]. <https://open.fda.gov/downloads/>. 2016.
- [10] FDA. About Device 510(k) [EB/OL]. [2016-10-12]. <https://open.fda.gov/device/510k/>.
- [11] 施雯慧, 陈颖, 姚捷, 等. FDA公共数据开放项目中屈螺酮炔雌醇片的分析研究[J]. 中国药物警戒, 2015, (9): 552-555.

(收稿日期 2016年11月4日 编辑 王雅雯)