



Mechanism Design of Health Information Platform Under the Big Data Environment

Ting Yu*, Xia Xue, Jiaqi Liu

Management School, Shanghai University of Engineering Science, Shanghai, China

Email address:

tingyu@sues.edu.cn (Ting Yu), M030215122@sues.edu.cn (Xia Xue), M030215108@sues.edu.cn (Jiaqi Liu)

*Corresponding author

To cite this article:

Ting Yu, Xia Xue, Jiaqi Liu. Mechanism Design of Health Information Platform Under the Big Data Environment. *Science Innovation*. Vol. 4, No. 2, 2016, pp. 105-109. doi: 10.11648/j.si.20160402.25

Received: April 6, 2016; Accepted: April 23, 2016; Published: May 5, 2016

Abstract: Focus on the current defects that the operator of health information platform in China hasn't built effective interfaces with the data source and information demand sides. By establishing the "3 + 2" platform conceptual model and designing the operating mechanism to break the bottleneck, the health data will be aggregated and used as much as possible. So that it will enhance the efficiency of the health information platform under the Big Data environment.

Keywords: Big Data, Health Information Platform, Defects, Mechanism Design

大数据环境下健康信息服务平台的机制设计

于挺*, 薛霞, 刘佳琪

管理学院, 上海工程技术大学, 上海, 中国

邮箱

tingyu@sues.edu.cn (于挺), M030215122@sues.edu.cn (薛霞), M030215108@sues.edu.cn (刘佳琪)

摘要: 针对当前中国健康信息服务平台运营主体未能与数据来源和信息需求主体建立有效接口这一显著缺陷, “3+2”平台的构建和有利于其运行与瓶颈突破的机制设计, 能促使健康数据最大限度地在平台汇聚与利用, 以期提升大数据环境下健康信息服务平台的效能。

关键词: 大数据, 健康信息服务平台, 缺陷, 机制设计

1. 引言

正式起步于1993年的信息化促使中国医疗卫生、医疗保险、药品生产与流通、医药教育与研发、医药监管等机构各自积累并不断产生了海量健康数据, 但这些数据之间的相互割裂使得健康信息服务平台难于在大数据环境下发挥促进公众健康的作用。2016年初, “百度疾病贴吧售卖”、“阿里药品监督码”等事件引起了社会各界的广泛

关注, 从根本上看则源于当前健康信息服务平台存在的缺陷。针对中国健康信息服务平台业已形成的数据来源丰富、信息需求强大、运营主体“弱、小、散”的“两头强中间弱”格局, 整合信息平台、俱乐部理论、网络外部性、健康经济等理论, 通过“3+2”平台架构并设计有利于其运行与瓶颈突破的机制, 以完善平台数据汇聚与有效利用的功能, 从而为“十三五”期间健康“大数据”实现“大价值”提供路径。

2. 文献回顾与问题的提出

2.1. 大数据引入前的健康信息服务平台研究

20世纪50年代中期,美国开始用计算机管理医院财务,是最早将信息技术用于健康领域的国家,1963年Tadao Umesao提出了信息化的概念[1],2001年 K Ono等人提出了信息平台的概念[2]。

对健康信息服务平台(HIS)的研究主要沿着技术应用与社会与经济两条相互交织的主线展开。从国际上看,各国学者对健康信息服务平台的研究从20世纪60年代初起步,最近三十年来,在技术应用层面主要集中于面向健康信息服务的信息技术应用与开发,如诊疗系统[3]、药品跟踪评价[4]等。在社会与经济层面主要涉及平台的法律、经济、伦理、社会问题、公共政策等多个维度[5-6]。

从中国国内看,学者们近二十年在技术应用层面的研究主要集中在平台与数据库构建[7]、数据交换接口[8]、信息共享[9]等方面。在社会与经济层面的研究则集中在国际范围内比较[10]、数据获取[11]等方面。

2.2. 大数据环境下健康信息服务平台研究

2008年《Nature》出版“大数据”专刊。麦肯锡认为大数据远超出常规软件工具获取、存储、管理、分析能力[12],大数据环境下在分析上追求相关关系而放弃对因果关系的渴求[13],海量数据冗余意味着信息需求主体所有分析目的都可能实现[14],并使海量数据高效采集、存储、分析成为现实[15]。

基于此背景,大数据环境下健康信息服务平台的数据汇聚、挖掘、利用已成为决定其效能发挥的核心。为此,在技术应用层面学者们主要关注面向健康这一特定目的的新技术平台架构[16]、数据存储[17]、数据挖掘[18]、数据隐私保护[19]等主题得到了密切关注,并已认识到平台在引入大数据后在疾病预防、精准医疗、新药研发、卫生控费等领域可能产生的潜在价值[20-21]。基于对数据汇聚障碍源于健康信息平台主体间相互割裂的认知,Robert G. Fichman等人[22]则从社会、经济、法律角度展开了讨论。

2.3. 大数据环境下各国相关政策与实践的描述

各国政府为推动健康数据的汇聚与利用积极提供政策支持。美国于2004年提出建立跨区域的国家卫生信息网,2015年国家卫生信息技术协调员办公室(ONC)发布报告《Connecting Health and Care for the Nation: A Shared Nationwide Interoperability Roadmap》。英国卫生部在构建NHS Choice信息平台的基础上,于2012年发布《信息的力量:让所有人获取所需要的卫生保健信息》。

在中国,医药卫生领域形成了丰富的信息平台,2009年在《中共中央国务院关于深化医药卫生体制改革的意见》的指导下,各地区纷纷搭建区域卫生信息平台;2015年国务院密集出台了11个关于促进大数据和“互联网+”的相关政策;国家人口与健康科学等数据共享公共平台建设得到快速推进;商业化平台方面,万达信息、卫宁软件等通过提供医疗信息化整体解决方案与系统维护,获得了

大量健康数据;以BAT为代表的信息服务企业涉足健康信息服务领域,在带来产业活力的同时也出现了“百度贴吧售卖”、“阿里健康药监码”等令人担忧的公共事件。

评述与前瞻:在长达数十年的信息化进程中,健康信息服务平台发展迅猛,但在大数据环境下,数据割裂、平台效能低下等问题日益突出。为此,学者们已在技术应用、社会与经济等层面对此症结进行了讨论,各国政府也尝试运用行政干预、资本力量加以突破。本文拟在前人研究基础上,选取社会经济视角对制约平台有效运行的问题进行探讨。

3. 当前健康信息服务平台的现状

健康数据是指有助于提升健康服务效率的电子化信息,如健康档案、诊疗病历、医药研发、制药与流通、医药卫生食品监管信息等。

3.1. 健康信息服务平台的主要类型

3.1.1. 按功能分类

根据各类主体承担的健康服务功能不同,可以把健康信息服务领域已经形成的平台划分为医疗服务、健康管理、医疗保险、药品生产、新药研发以及医药卫生监管等类型。

3.1.2. 按性质分类

根据平台运营主体的所有制性质,可以将现有平台划分为公共性与商业化平台两类,前者如卫生部门建立的各类区域卫生信息平台、药品监督部门组建的综合业务平台,后者如阿里健康、万达信息等。

3.1.3. 按生命周期不同阶段分类

针对居民从孕育、出生、成长到衰老等不同生命周期,健康领域都有不同机构提供相应的服务,并由此形成了各类信息服务平台,如妇幼卫生保健平台、针对学生身体素质的信息管理系统,服务职场人士的体检信息平台等。

3.2. 当前健康信息服务平台的构成与现状评价

当前,健康信息服务平台由数据来源主体、平台运营主体、信息需求主体三部分构成,如图1所示。

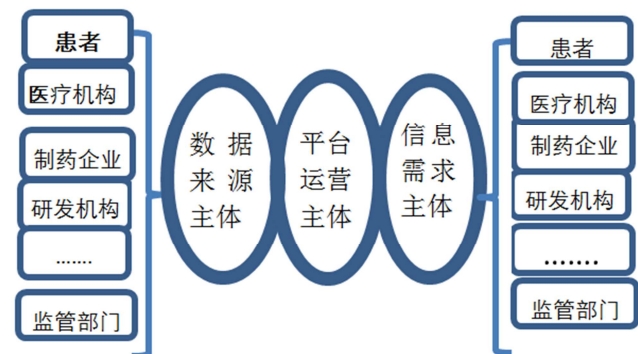


图1 当前健康信息服务平台的构成情况。

从平台的参与主体来看,患者、医疗机构、制药企业、医药研发、监管部门等既是平台的数据来源主体也是信息需求主体,两类主体的重合度高;从数据权属来看,健康数据分别来自居民、医疗机构、研发机构等多重主体,由健康信息平台采集汇聚、存储、分析、发布,这一过程十分漫长,其权属关系相当复杂,数据权属分割体系尚未有效确立[23];从平台的信息流动来看,由于平台之间缺乏有效的信息共享机制,平台信息流呈现短链、并行、割裂的“链状”结构。

4. 当前健康信息服务平台的缺陷与瓶颈分析

4.1. 当前平台存在的缺陷

与其它类型信息服务平台的运营主体强而数据来源与使用两侧弱不同,健康信息服务平台则是数据来源与使用两侧强而运营主体则呈现弱、小、杂的态势。从数据来源侧看,健康数据主要由各类机构拥有,由于存在数据格式、数据权属、数据成本与收益、平台未来发展方向不一致等问题,各类机构的数据难以汇聚;从信息需求侧看,当前各类健康数据拥有者各自持有、使用自有数据,并各自形成了内部规章,这使得平台既难以有效导入海量数据,也无法有效满足各类主体强烈的信息需求。换言之,当前平台运营主体未能与数据来源侧和信息需求侧建立有效接口。

4.2. 引发缺陷的瓶颈分析

4.2.1. 存在健康隐私泄露的风险

David C. Vladeck曾举例用户在网上搜寻油炸锅的例子,作为数据挖掘者可以追踪到用户点击过的所有页面,进而解读其具有不健康的饮食习惯,而以这些数据为基准的信息很可能会到达健康顾问或是潜在雇主的手上[24]。健康信息服务平台虽承担着对个人信息“脱敏”的责任,但大量数据的流出仍可让已被去除个人隐私的数据通过相关性分析而与居民个人建立联系,从而导致其健康隐私泄露。

4.2.2. 数据权属难以合理分割

以诊疗过程为例,健康数据在形成过程中往往涉及居民、医疗机构、医疗保险、研发机构等多重主体。在社会保险支付占主导地位的医疗体制下,居民就诊时若要获得保险赔付就必须使用医保卡、采用实名制,这时居民其实已经将其健康数据权属以默认转移机制部分地让渡给了医生、医院、社会保险机构等主体,而健康数据一旦形成又必然存在多个用途,且复制的成本极低。由于各类主体在健康数据权属上缺乏可行的分割方法,这将阻碍数据的导入与进一步的应用。

4.2.3. 隐性矛盾显性化

健康信息服务平台对健康数据大范围的汇聚与深度挖掘将使误诊、药物滥用等易于界定,这会使隐含在数据中的社会矛盾显性化,从而加剧医患矛盾。

4.2.4. 平台运营不可持续

一方面数据在使用中会产生收益,另一方面,数据在平台上的存储、调用、分析则会形成成本,然而这些收益和成本在质和量上是不匹配的,致使平台难以实现持续运营。这也很好地解释了一些商业化健康信息服务平台时常出现过度商业化行为的现象。

4.2.5. 监管风险凸显

以BAT为代表的非健康领域的商业化信息服务平台转而从事健康信息服务的实践探索,在打破公共健康信息服务平台垄断的同时也带来了公众对其公信力、盈利合理性等方面的质疑,致使监管风险凸显。

5. 健康信息服务平台“3+2”架构与运行机制设计

5.1. 平台的“3+2”架构

针对上述缺陷与瓶颈,本研究拟在当前平台“数据来源-平台运营-信息需求”三个构成部分基础上,突出数据来源与信息需求两侧接口的地位,形成健康信息服务平台“3+2”概念模型(如图2所示)。

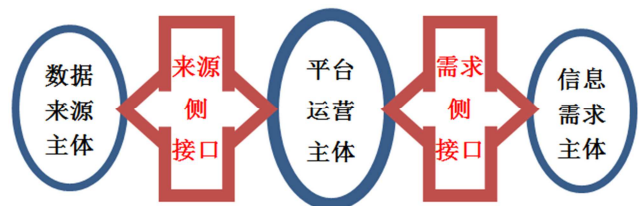


图2 “3+2”的健康信息服务平台概念模型。

来源侧接口的任务是把数据来源主体与平台衔接起来,使健康领域内各类主体产生和积累的数据充分地导入平台;使用侧接口则担负着衔接平台与各类健康信息需求主体的任务。

5.2. 完善两侧接口功能的运行机制设计

5.2.1. 建立两侧接口与平台的衔接机制

通过设计物理接口,让平台能够兼容各种格式的数据,同时形成一套协议,以数据兼容、数据权属清晰、利益相容为原则,构建数据采集、存储、分析、调用、监管所必要的行为机制,让健康数据最大限度地平台汇聚,并使数据得到合规、高效利用。

5.2.2. 设计两类信息共享机制

平台的数据来源与信息需求主体高度重合,因此,共享机制是平台有效运行的重要机制。当使用主体需要与其提供的是同类信息时,适用于同质信息自返共享机制;当使用主体需要与其提供的是非同类信息时,则适用于异质信息交叉共享机制。只有两类信息共享机制共同起作用,才能实现数据的“冗余”与多源异构,进而提高数据分析的可靠性。

(1) 同质信息自返共享机制

与其它信息服务领域有所不同，健康数据的提供者往往也是同类数据的需求者。如百度血友病贴吧中的患者在分享自身数据的同时也期望共享病友或专业人士提供与其所患疾病相关的信息。由于同质信息存在频繁、连续的交换，如果每一次数据的交换都采取交易购买的方式，势必导致数据使用成本急剧上升。结合俱乐部理论，在保证数据权属清晰、健康隐私得到保护的前提下，对于同类健康数据的共享可建立“俱乐部会员式”的同质信息自返共享机制，即主体通过提供数据而获得分享其它同类主体拥有数据的权力。

(2) 异质交叉信息共享机制

对于异质健康数据，由于主体之间数据交换的频率较低，因而，可以在数据权属清晰、健康隐私得到保护的前提下建立起数据交易机制，即在获取非同类主体数据时需要进行经济支付。例如，通常来说研发机构的实验数据医院是不需要的，而研发机构对医疗机构的数据也并不总是感兴趣，但是在研发需求确立阶段，的确需要医疗机构能够提供患者针对某类疾病的用药效果，这时，研发机构就应该向医院付费获得相应数据。

当然，以数据互换为基础的同质信息自返共享机制并不排斥以交易为核心的异质交叉信息共享机制，毕竟一些加工后的数据与初始数据的价值不可同日而语；同样，适用异质交叉信息共享机制的相关机构也有机会获得业内的免费数据。而且，随着健康信息服务平台促进健康、引导舆论、维护公众健康权益的职责进一步落实，越来越多的平台将免费开放更多的健康数据。

5.2.3. 健康信息服务平台的运营与监管

一是应明确平台运营主体的资质。主体应在数据采集、存储、分析，辨识需求主体类别，确保信息安全等方面能力突出，可以是具备资质的政府所辖行政事业单位、国有企业或商业机构。

二是需要设计平台运营机制。平台运营特别要注意保护公众健康权益，逐步扩大普遍服务范围。因此，需要设计可实现公众、机构、商业资本、行政等主体诉求的规则。其运营机制如图3所示。三是构建包括平台自律、利益相关者相互约束、行政监管三个部分的平台监管体系。

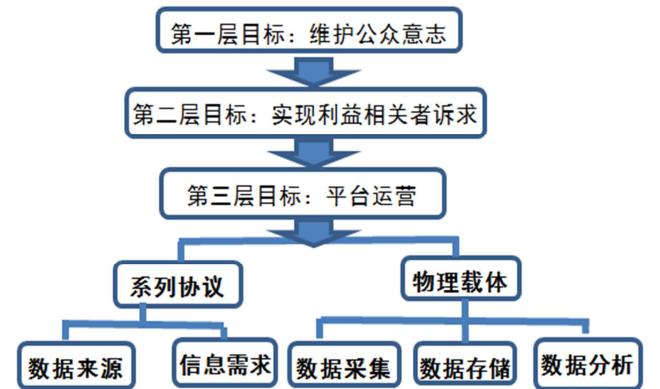


图3 大数据环境下健康信息服务平台的运营机制。

6. 健康信息服务平台未来发展展望

6.1. 平台运营主体展望

由于网络外部性的存在，健康信息服务平台的总体数量将逐步减少，盈利性与非盈利性平台保持共存，平台主体将得到严格监管。

6.2. 平台信息流将由“链状”向“环状”结构演变

两侧接口功能实现后，数据将由分散、孤立、断裂变得整合、全面、连续，而强大、频繁的信息需求将打破来源侧数据既有结构，促进平台信息流向“环状”结构演变（如图4所示），使健康“小数据”转为“大数据”。

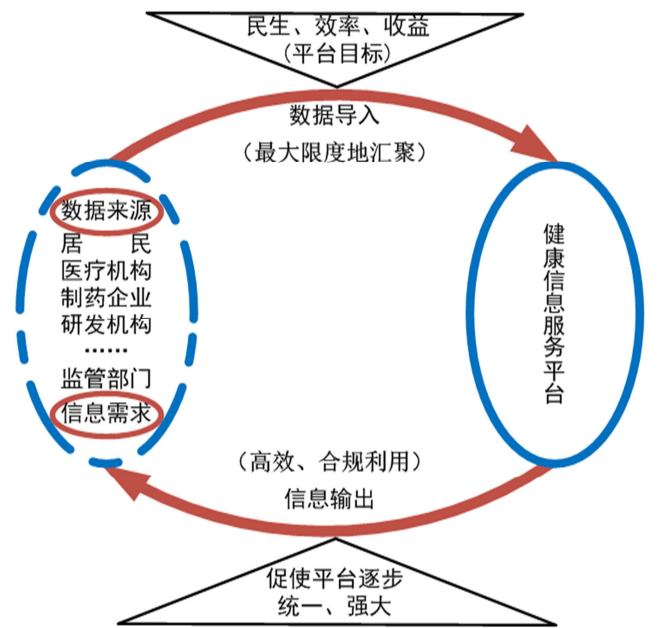


图4 平台信息流“环状”结构示意图。

6.3. 平台效能展望

大数据环境和平台的环状信息流，将使以相关分析为基础的推断技术获得巨大的应用空间，与以因果分析为基础的证明技术互补，可以整体提升健康服务业的效能。

7. 结语

目前中国健康信息服务平台面临最大的缺陷在于健康信息服务平台运营主体未能与数据来源和信息需求主体建立有效接口，导致这一问题的根源在于健康信息服务平台面临健康隐私泄露的风险、数据权属难以合理分割、隐性矛盾显性化、运营不可持续、监管风险凸显这五大瓶颈。本研究通过构建“3+2”的健康信息服务平台，在数据来源侧与运营平台之间，以及运营平台与数据需求侧之间建立两个接口，并建立信息共享机制，从而充分发挥健康信息这一资源的价值，提升健康产业的服务水平。

参考文献

- [1] Tadao Umesao. Information industry theory: Dawn of the coming era of the ectodermal industry [J]. Hoso Asahi, 1963.
- [2] K Ono, K Maruyama. Information platform: Concepts and research topics [J]. Nii Journal, 2001, 3(3): 1-5.
- [3] CG Schroeder, PG Pierpaoli. Direct order entry by physicians in a computerized hospital information system [J]. American Journal of Hospital Pharmacy, 1986, 43(2): 355-359.
- [4] RA Raschke, B, Gollihare, et al. A computer alert system to prevent injury from adverse drug events: development and evaluation in a community teaching hospital [J]. The journal of the American Medical Association, 1998, 280(15): 1317-1320.
- [5] Mike W. Chiasson, Elizabeth Davidson. Pushing the contextual envelope: developing and diffusing IS theory for health information systems research [J]. Information and Organization, 2004, (14): 155-188.
- [6] JG Anderson. Social, ethical and legal barriers to e-health [J]. International Journal of Medical Informatics, 2007, 76(76): 480-483.
- [7] 陈戏墨, 林超华, 等. 医院多系统异构多应用软件集成平台的研究[J]. 软件导刊, 2011, (3): 12-14.
- [8] 蒋巍巍, 王海舜, 等. 面向健康体检信息共享的HL7接口引擎设计[J]. 医学信息, 2009, 2(9): 1701-1703.
- [9] 沈剑峰, 汪巍等. 基于卫生信息平台的电子病历信息共享研究和实现 [J]. 中国生物医学工程学报, 2013, 32(4): 504-507.
- [10] 沈丽宁. 国外健康信息服务现状扫描及启示[J]. 医学信息学杂志, 2010, 1(6): 38-40.
- [11] 姚维保. 公共健康信息的公共获取问题研究[J]. 图书情报工作, 2004, 48(5): 10-12.
- [12] McKinsey Global Institute. Big Data: The next frontier for innovation, competition, and productivity[R]. Technical report, 2011.
- [13] 维克托·迈尔·舍恩伯格. 大数据时代[M]. 杭州: 浙江人民出版社, 2012.
- [14] 麦尔荀伯格·库基耶著, 林俊宏译. 大数据[M]. 台北: 远见天下文化, 2013.
- [15] 涂子沛. 大数据[M]. 林: 广西师范大学出版社, 2013.
- [16] Q Yao, X Han, et al. Cloud-Based Hospital Information System as a Service for Grassroots Healthcare Institutions [J]. Journal of Medical Systems, 2014, 38(9): 1-7.
- [17] 杨剑, 胡新平, 等. 基于云的区域电子健康档案存储和交换架构[J] 中国数字医学, 2010, 1(3): 10-13.
- [18] CP Shen, C Jigjidsuren, et al. A Data-Mining Framework for Transnational Healthcare System [J]. Journal of Medical Systems, 2011, 36(4): 2565-2575.
- [19] 陈鹤群. 大数据环境下医疗数据隐私保护面临的挑战及相关技术梳理[J]. 电子技术与软件工程, 014, (16): 51-53.
- [20] TB Murdoch, AS Detsky. The Inevitable Application of Big Data to Health Care [J]. The journal of the American Medical Association, 2013, (13): 1351-1352.
- [21] 王潇, 张爱迪, 等. 大数据在医疗卫生中的应用前景[J]. 中国全科医学, 2015, 16(1): 113-115.
- [22] Robert G. Fichman, Rajiv Kohli, et al. The Role of Information Systems in Healthcare: Current Research and Future Trends [J]. Information Systems Research, 2011, 22(3): 419 - 428.
- [23] MA Rodwin. The case for public ownership of patient data [J]. Jama the Journal of the American Medical Association 2009, 302(1): 86-88.
- [24] David C. Vladeck. Information Access-Surveying the Current Legal Landscape of Federal Right-to-Know Laws [J]. Texas Law Review, 2008, 86(7): 1787-1836.