

数据科技:智慧金融的技术基础

朱扬勇¹,胡乃静²

(1.复旦大学,上海200433; 2.上海金融学院,上海201209)

摘 要:数据科技是开发网络空间(CYBER SPACE)数据资源所用到的科技,是发展包括智慧金融在内的智慧产业的技术基础。本文介绍了数据科技和数据产业的内涵,探讨了数据在各类智慧系统中的核心作用,以数据科技为技术基础的智慧金融,提出了发展金融数据共享规范、金融数据产品、绿色存储、高可信环境等建议,通过发展数据科技推进智慧金融的进步。

关键词:数据科技;智慧金融;金融数据

中图分类号:F830.49

文献标识码:A

文章编号:1673-680X(2012)01-0016-07

一、数据资源是战略资源

国民经济与社会发展的信息化是国家战略的重要组成。信息化的本质就是将现实世界中的事物以数据的形式存储到计算机系统中,即信息化是一个生产数据的过程。随着信息化进程的推进和发展,国家和社会都是建立在计算机系统之上的,整个社会将依靠计算机软件在运转,运转将根据数据来进行,运转地结果将以数据的形式记录在计算机系统中。这样,数据被大量生产并存储在网络空间而形成数据资源。由于这些数据是人们工作、生活和社会发展的记录,并被存储在计算机数据库中,从而形成了海量的数据资源,包括:金融经济数据、社会行为数据以及更大量的互联网行为(电子商务行为、网络游戏行为、电子邮件行为、网络社区)等等。

由于互联网的发展,这些数据资源几乎无时不刻的在产生:发送的邮件、工作产生的文档、网络上传的照片、银行产生的交易记录、甚至政府机构收集

收稿日期:2011-12-15

作者简介:朱扬勇(1963-),男,浙江武义县人,复旦大学计算机科学技术学院教授,博士生导师,数据科学研究中心主任。

胡乃静(1971-),男,山东栖霞市人,上海金融学院信息管理学院教授。

和管理的社会数据等。IDC 曾经对全球的数据增长状况做过调查,显示在 2008 年这一年内,就产生了约 4870 亿 GB 的数据,如果将这些数据出版成书排成一列,能在地球和冥王星之间来回 10 次。而随着互联网应用的深度发展,数据更将以指数级继续增长。

数据持续的爆炸式增长及其极大的开发利用价值,数据资源已经成为重要的现代资源。数据资源在本世纪有可能超过石油、煤炭、矿产,成为最重要的人类资源之一。这是因为社会、政治和经济的运转都将依赖于数据资源,而石油、煤炭、矿产等资源的勘探、开采、运输、加工、产品销售等等无不依赖数据资源,离开了数据资源,这些工作都将无法开展。所以,对数据资源进行保护和开发利用所形成的数据产业,将会成为新的支柱产业。目前已经具有代表性企业有 Google、百度、万得、万方、大智慧等,特别注意,Google 可能已经是全球最大的数据资源中心,在可预计的未来,Google 可能成为全球企业之首。

二、数据产业及其发展

1. 数据产业

数据产业是对网络空间数据资源开发利用所形成的产业,其产业链从网络空间获取数据并进行整合、加工和生产,然后到数据产品的传播、流通和交易,以及相关的法律和其他咨询服务等,内容涉及数据资源勘探、数据资源整合、数据资源管理、数据资源权益和共享、数据资源开发、数据资源服务等方面的经济活动。目前,数据产业涵盖了数字出版与文化业、电子图书馆和情报业、多媒体产业、数字内容业、数据服务业和信息咨询业、领域数据资源开发服务业(政务、商务、科学、社会、金融、经济、地理等等)等多个产业。

“数据产业”与“信息产业”相互联系,又有所不同。信息产业涵盖几乎所有电子技术的所有环节,包括集成电路、电子元器件、电气产品、电子数字产品、智能设备、计算机及其软件、计算机网络等等。信息产业发展的结果是在计算机系统中积累了庞大的数据资源,因而数据产业可以认为是信息产业的一部分。但另一方面,由于目前数据资源已经非常庞大,足以与石油煤炭资源相媲美,无论从技术还是从产业上都需要从根本上关注数据,这使得单独发展数据产业很有必要,因此需要将数据产业从信息产业中分离出来,快速推进。

事实上,近年来数据产业作为一个独立的产业已经得到了极大的发展。相比于信息产业的部分产品发展速度,数据产业已飞速发展。

2. 数据产业的特点及其发展

当前,世界各国都在利用卫星、遥感、飞机、勘探等手段收集宇宙、大气、地球、海洋等自然数据,形成自然数据资源;利用 DNA 测序获得关于生命的数据,形成生命数据资源;而国民经济与社会信息化则产生了社会发展和人类行为的数据,形成了经济社会数据资源。

首先,掌握数据资源将在国际上掌握主动权。不论是反倾销诉讼、铁矿石谈判、汇率问题、节能减排、碳关税谈判等等重大国际政治经济事务,无一不是靠数据说话的。相关数据在计算机系统中是存在的,需要将它们开发出来,为我们的政治经济发展服务。

其次,掌握数据科学技术才能开发利用数据资源。这就需要支持数据科学技术研究,发展数据产业。发展数据产业本身就可以产生巨大的经济和社会效益。甚至可以说,如果在数据科学技术方面领先,那么在现代军事方面都有可能处于领先地位(例如,美国成立了网络部队司令部,这样的司令部在网络环境下靠的是数据资源和数据科学技术)。

3. 数据产业为政府、企业和公众服务

数据产业显然是为政府和公众服务的现代服务业。随着当前数据仓库数据挖掘等数据科学技术的应用,这些政府社会和企业数据资源开发做了一些工作,但离数据产业发展后的效果还差很远。

政府方面,在电子政务发展多年后,政府的数据资源是最丰富的,其数据资源的开发利用将提升政府决策的科学水平。事实上中共中央办公厅和国务院办公厅早在2004年就曾经发文要求加强信息资源开发利用。企业方面,企业的选址、新产品开发、市场定位和市场营销、售后服务等都是建立在行业和企业数据资源上的。社会公众方面,Google、Baidu、大智慧、钱隆是我们所熟悉的数据产业企业,没有这些企业,很多工作都将回到原始状态。

4. 数据产业是新经济的源泉

新经济是以高科技、信息、网络、知识为其重要构成部分和主要增长动力的经济,它最早源于美国。1990年代以来,美国社会经济一直呈现高增长、低通胀、低失业的特点,经济界把这种现象称为“新经济现象”。新经济是跨越地理的经济,需要区域协作,形成区域联合体。新经济以高科技、信息技术等为增长原动力,以电子商务为主要交换手段,趋向全球一体化。发展新经济要做到技术、观念、市场、制度四方面创新。新经济的典型企业,国外有:YAHOO, GOOGLE, eBay, amazon, SECONDLIFE等等;国内有:百度,新浪,网易,阿里巴巴等等。

显然,主要的新经济企业都是建立在数据资源之上的,以数据资源开发利用为主要工作的。经济结构调整的一个重要任务是大力发展新经济,在开发自然数据资源、生命数据资源、各行各业数据资源过程中将产生众多新经济企业。这是值得我们政府去推动的。

三、数据科技是智慧金融发展的技术基础

1. 数据是智慧系统的基础

数据科技是研究网络空间数据界(DATANATURE)的现象和规律的科学,数据科技被应用于网络空间中数据资源的开发利用。

目前,数据资源开发利用问题存在着大量技术问题有待解决。各数据资源中心如何共享、如何消除数据之间的不一致、如何分析数据的规律等等仍然没有有效的解决方法;国家在打击网络不良信息、管理舆情民情、促进和谐社会建设方面,还没有很好的科学技术手段;而电子商务交易、网络游戏、电子出版物和作品的权益等方面需要先进的数据科技。

“智慧地球”概念的发展,进一步引发了“智慧金融”、“智慧电力”、“智慧医疗”、“智慧城市”、“智慧交通”、“智慧供应链”等领域的不断“智慧化”,形成了各个领域的智慧化系统。但从数据科技的角度来看,目前智慧系统的状况是:大部分 BI 系统不具有实用的数据挖掘能力;大部分系统数据收集不全,且数据不一致;大部分系统数据增量更新有问题;大部分 BI 系统只提供了报表能力,不具备智慧能力。

一个“智慧”的系统应该满足三方面的要求:(1)对未来问题的预测或预判;(2)对已有问题的总结;(3)异常问题的应急处理。这些取决于一个完整的大规模数据资源中心和良好的数据分析软件,即取决于数据科技的掌握和应用。因此,一个智慧的系统应当以数据及其处理为核心,才能够发挥其智慧的优势。

2. 以数据科技为核心的智慧系统架构

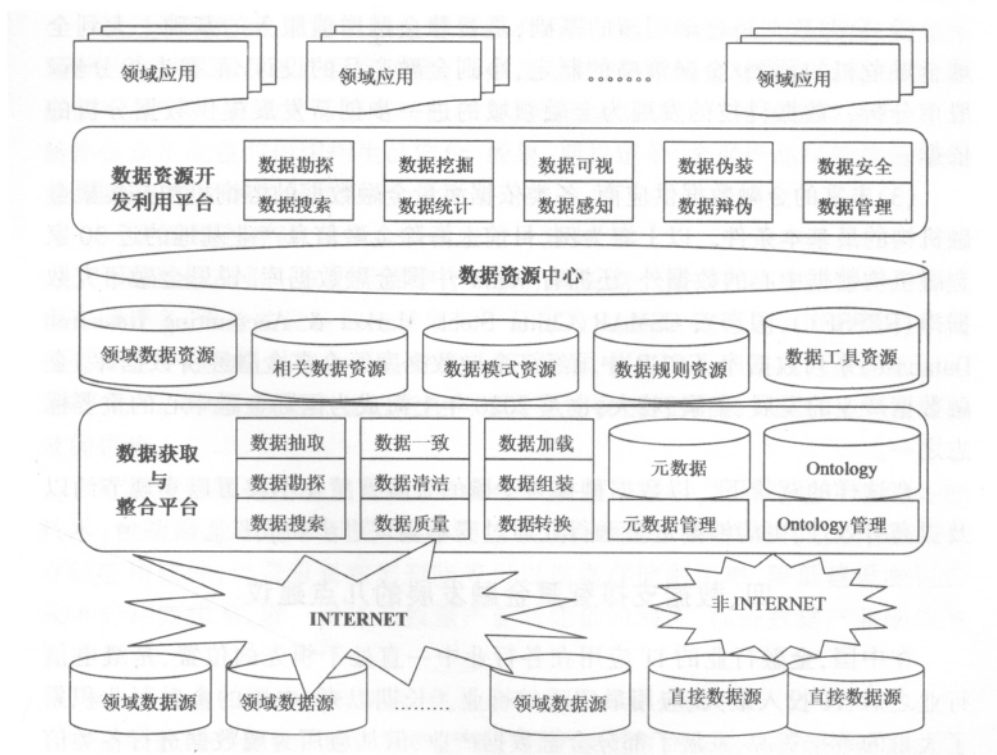


图1 智慧系统的核心框架图

如图 1 所示,一个智慧系统的核心应当包括一个数据获取与整合平台,能够从网络空间获取需要的数据,并能够整合成一个数据资源中心,数据资源中心应该是完备的、可用的和最新的,同时应当有一个数据资源开发利用平台,提供各种数据分析工具,以实现“智慧”。

3. 数据科技支撑智慧金融的发展

智慧金融主要通过信息技术的支撑使现代金融行业在组织结构、业务流程、业务开拓以及客户服务等方面得到全面提升。IBM 认为智慧金融在金融的业务和技术层面具有更透彻的感应度量、更全面的互联互通、更深入的智能洞察三大特征,其构建贯穿金融业务的前中后台,通过新兴技术和工具来实现、发展。

从数据的角度来看,现代的金融已融入了更多的数字信息元素,智慧金融的发展更加依赖于数据的支撑:

(1) 处于高度信息应用的金融领域,数据及数据处理是其业务运行的基础,脱离了原始数据的支撑,任何金融业务将无法开展。目前,金融业的各个应用领域正在发展各自的数据科技:银行业通过数据大集中工程,各自形成了一至多个行内的数据中心;交易所、证券业也已形成了各自独立的数据中心并对外出售数据;保险业也开始逐步建立起自己的数据中心。

(2) 金融数据是金融创新的基础,是智慧金融增值服务的基础。大到全球金融危机的预测/金融策略的制定,小到金融产品的设计/汇率走势分析/股市分析,数据科技的发展为金融领域的进一步创新发展提供数据分析的依据。

(3) 丰富的金融数据供应商、各类依据海量金融数据的咨询机构是集聚金融机构的最基本条件。以上海为例,目前上海除金融信息产业基地的近 30 家金融机构数据中心的数据外,还拥有 Wind 中国金融数据库、锐思金融研究数据库(RESET),国泰安 CSMAR(China Stock Market & Accounting Research Database)系列数据库,CCER 中国经济金融数据库等众多金融经济数据库。金融数据产业的发展、集聚、壮大,也是 2020 年上海成为国际金融中心的重要标志之一。

在这样的背景下,以数据科技为支撑的金融智慧化不仅可以实现节约以及提高有效性,更能够推进金融行业通过变革实现进步和成长。

四、数据支撑智慧金融发展的几点建议

在中国,金融行业的 IT 应用在各行业中一直处于领先的位置,是继电信行业之后,IT 投入最大、应用最广泛的行业。长期以来,中国的金融行业积累了大量的数据资源,发展了部分金融数据产业,但从应用金融数据进行各类信息系统的服务、金融创新的产品开发、金融风险的监控等智慧金融的应用发展

来看,相比于纽约、伦敦等国际金融中心,中国的金融业明显存在应用层次较低、范围不广、深度不够等问题。因此,提升数据及其产业,才能更好地支撑智慧金融的发展。

1. 建设各类金融数据共享规范,推进金融数据产业发展

中国的金融数据产业虽然有所发展,但大部分规模较小,产值较小。2009年全球金融数据 226.8 亿美元的市场中,中国的金融数据市场支出仅 3.42 亿美元,并且存在着分散、重复等问题。

建设各金融数据中心的数据共享规范,形成整合各类公开数据、中小金融资讯机构的单项数据等资源的基础。可以打破金融数据中心各自建设、各自运营的模式,产生广泛的来源获取、量度、建模、处理、分析大量结构化和非结构化金融数据的能力,形成延伸至金融系统各个相关领域的信息收集模式,从而推进金融数据产业的发展,提升金融机构、金融企业等智慧金融产业的发展环境。

2. 通过技术发展金融数据产品,提升金融数据的增值服务

中国虽然已经建立了各类金融数据中心,但金融数据产业在宏观数据分析、复杂数据分析方面尚未真正起动,对金融数据的后续加工和增值服务欠缺,存在着“原料”型服务的弊端。

通过数据挖掘、数学模型分析等技术手段,促进以金融数据分析与处理为基础的金融数据服务产品的发展,形成金融数据分析、金融资产风险管理、金融衍生品定价、自动交易策略模拟和交易模型验证评估等增值服务能力,为市场提供固定收益与信用衍生品定价、股票、期权定价、金融产品风险度量等服务,推进智慧金融的发展。

3. 发展绿色存储,降低金融数据产业能耗

金融数据产业一方面需要整合大量计算资源,另一方面也需要大量数据库存储设备,而这些设备的能量消耗十分惊人。Google 每建立一个数据中心都需要建立专门的电站来为设备提供能源,目前,信息产业已经成为全球第五大能耗大户。因此,数据产业的能耗问题需要重视,否则将极大限制其应用和普及的进程。

针对金融数据产业的特点,建设金融数据云中心,发展绿色节能的各种技术,包括磁盘和磁带系统、服务器连接、存储设备、网络架构、文件服务和存储应用软件,以及重复数据删除等可以提高存储利用率、降低建设和运行成本的存储技术,进一步降低数据产业对能量的消耗,推进数据产业的应用发展。

4. 建设高可信的执行环境,提高数据产业的安全性

金融数据产业意味着可能要把客户信息这类很高商业价值的数据存放到计算服务提供商的手中,然而,一方面服务商希望各种服务可以持续,从而保

证服务质量,另一方面,随着数据集成度越来越高,数据的安全性也越来越重要,因为不论从用户角度,还是从国家安全角度,哪怕是数据碎片的泄露也是危险的。

进一步发展数据安全技术,建设高可信的数据产业的运行环境,避免物理资源共享所带来的数据安全危机,可以为各种应用提供稳定的服务环境。另一方面,开发各种隐私安全保护技术,避免在数据服务中泄露各类隐私信息,可以进一步保障数据产业服务的安全性,提升数据产业的发展。

参考文献:

- [1] 李庆莉.金融数据中心建设:迈向国际一流[J].中国金融电脑,2011,(12).
- [2] 胡乃静.上海的智慧金融及其发展探讨[J].上海金融学院学报,2011,(1).
- [3] 付小龙,袁芳,张民社等.国家自然科学基金委员会数据资源规划研究与实践[J].中国科学基金,2011,(05).
- [4] 张桂刚,李超,张勇等.云环境下海量数据资源管理框架[J].系统工程理论与实践,2011,(S2).
- [5] 赵永耀,秦志光.基于客户需求的移动数据产业生态系统研究[J].管理学报,2009,(09).

Data Technology: Wisdom Financial Technology Foundation

Zhu Yangyong, HuNaijing

Abstract: Data technology is the science and technology used by the development of network space (CYBER SPACE) data resources, is the development basics of financial market, including wisdom wisdom technology industries. This paper introduces the data technology and data industry connotation, discussed the data in all kinds of intelligent system in the central role, and data technology based on wisdom finance, put forward to develop the financial data, financial data sharing standard products, green storage, high trustworthy the proposal such as the environment, through the development of Science and technology to promote the progress of the wisdom of financial data.

Key words: Science and technology intelligence data; finance; financial data

(责任编辑:顾广贤)